Psychological Bulletin
1995, Vol. 118, No. 3, 422–429

0033-2909/95/$3.00

Nonlinear Models of Clinical Judgment: Meehl's Data Revisited

Yoav Ganzach
Tel Aviv University

Previous attempts to detect nonlinearity in clinical judgments have not succeeded because of a lack of good nonlinear models. Much research in this area was based on data collected by Paul Meehl, which include clinicians' judgments of mental disorder on the basis of Minnesota Multiphasic Personality Inventory profiles. In this article, Meehl's data are reanalyzed using several versions of the scatter model in which nonlinearity is represented by the within profile scatter(s) of the cues. The author finds that these versions give a better fit to the data than the linear model. He also finds systematic patterns of nonlinearity that lend themselves to psychological interpretation.

Are clinical judgments nonlinear? That is, are the rules governing these judgments different from rules in which judgment is a linear combination of cue values? Although both process tracing studies (e.g., E. J. Johnson, 1988; P. E. Johnson, Hassebrock, Duran, & Moller, 1982) and clinician intuition about their judgments (e.g., Meehl, 1954) suggest that clinical judgments are nonlinear, statistical analyses of these judgments (see Hoffman, 1960, for a pioneering treatment of such analyses; see Brehmer & Brehmer, 1988, for a recent review) and, in particular analyses of judgments that are representative of the environment, have generally not found such nonlinearity (Slovic & Lichtenstein, 1971). Linear regression models usually give at least as good a fit to clinical judgment as nonlinear regression models which are designed to capture the nonlinearity in the judgment process.

The fact that linear models give a good fit to clinical judgments does not mean, of course, that the judgment process is linear. It is well known that linear models give a good fit to judgment even if the underlying process is highly nonlinear (e.g., Dawes & Corrigan, 1974). However, the lack of regression models that are capable of capturing nonlinearity in the judgment process is troublesome because it makes the study of clinical judgment all the more difficult.[1]

Two of the most well-known studies that led to the pessimistic view about the ability to model nonlinearity in clinical judgment were based on data collected by Paul Meehl in the 1950s. These data included evaluations of 861 patients, diagnosed as either neurotic or psychotic on the basis of their Minnesota Multiphasic Personality Inventory (MMPI; Meehl, 1959) profiles—their scores on eight clinical scales and three validity scales of the MMPI. The evaluations were made on an 11-step forced normal distribution from least psychotic (1) to most psychotic (11). They were obtained from 13 clinical psychologists and 16 clinical psychology trainees (see Meehl, 1959, for a detailed description of the data).

In one of these two studies, Wiggins and Hoffman (1968) examined the fit of two nonlinear models: the quadratic model, which included, in addition to the linear components, 66 terms composed of the squares and products of the cues; and the sign model, which included, in addition to the linear components, 59 signs (nonlinear combinations of scales). The authors' findings did not show any consistent differences between the linear model and the nonlinear models. In the second study, Goldberg (1971) examined four nonlinear models, including Einhorn's (1970) hyperbolic and parabolic models, and found that, for 26 out of the 29 judges, the linear model gave the best fit for the data.[2]

That this previous work did not detect consistent reliance on nonlinear strategies is particularly troubling because there are indications that nonlinear elements do exist in the clinicians' judgments. For example, Goldberg (1971) reported that a number of nonlinear signs had a significant correlation with the residuals (from a linear model) of the judgments. The problem is that these indications for nonlinearity do not constitute a meaningful and generalizable pattern that allow for the development of a general model for nonlinearity in clinical judgment.

One reason for the previous failures in modeling nonlinearity in Meehl's data is the models' lack of statistical power (see Ganzach & Czaczkes, in press, for further discussion of this issue). The lack of statistical power is especially problematic because, as is typical for representative data, the cues of MMPI profiles are highly correlated (high-intercue correlations decrease the difference between the fit of various alternative models). Another reason for these failures may be an incorrect specification of the nonlinear processes underlying judgment. Thus, for ex-

I thank Lewis R. Goldberg for his generous support both in supplying the data and in answering many of my questions. This article, however, does not necessarily reflect his views. I also thank Benjamin Czaczkes and Danny Mingelgrin for their help in all parts of this research.

Correspondence concerning this article should be addressed to Yoav Ganzach, School of Business Administration, Tel-Aviv University, Tel-Aviv 69978, Israel.

[1] Note that models other than regression models were also used in studying nonlinearity. In particular, many computer models represent judgment and decision as nonlinear process (e.g., Ben-David & Mandel, in press; Kleinmuntz, 1963).

[2] In addition to studies that attempted to examine for nonlinearity in judgment, there were a number of studies that attempted to examine for nonlinearity in the criterion (e.g., Meehl, 1959; Meehl & Dahlstrom, 1960). The focus of the current article is, however, solely on nonlinearity in judgment.

ample, the nonlinear process in Wiggins and Hoffman's (1968) quadratic model presupposes nonlinear relationships between all pairs of the MMPI scales, which is highly unlikely given people's limited processing capacities.

Although the modeling of nonlinearity in clinical judgments that are based on representative stimuli was rather disappointing, there was somewhat more success in the modeling of judgments elicited in laboratory settings and based on "artificial" stimuli (e.g., stimuli in which cues are orthogonal or even negatively correlated). Early examples are Einhorn (1971, 1972); Einhorn, Komorita, and Rosen (1972); and Wright (1974). These studies used Einhorn's hyperbolic and parabolic models. In later studies, Brannick and Brannick (1989) and Ganzach (1993, 1994b) used a new model labeled the *scatter model* to detect nonlinearity in judgment. Recently, Ganzach and Czaczkes (in press) showed that this model outperformed Einhorn's models in a number of data sets obtained from laboratory experiments. Because the detection of nonlinearity in Meehl's data appears to be the benchmark for any model aimed at detecting nonlinearity in representative judgment in general, and in clinical judgment in particular, this article shows results concerning the ability of the scatter model to detect nonlinearity in Meehl's data. In addition, the article develops this model beyond the simple version that was used in Brannick and Brannick's and Ganzach's studies by introducing versions of this model that better capture nonlinearity in nonorthogonal, real-life stimuli.

Disjunctive Strategy, Conjunctive Strategy, and Scatter

Two important types of nonlinear strategies are the *disjunctive* strategy and the *conjunctive* strategy (e.g., Dawes, 1964). In these strategies, the impact of a cue depends on its rank relative to the other cues. In the disjunctive strategy, judgment is based primarily on the higher cue(s). In the conjunctive strategy, judgment is based primarily on the lower cue(s).[3] Note that high–low cues do not necessarily reflect a value judgment (i.e., they do not mean favorable or unfavorable) but are defined vis-à-vis the judgment. The convention used in this article, which is implicitly implied in previous articles (e.g., Einhorn, 1971; Goldberg, 1965, 1971; Ogilvie & Schmitt, 1979), is that cue values are scaled to have a positive correlation with judgments.[4]

Disjunctive and conjunctive strategies are associated with the scatter, or gap, between cue values. To illustrate, consider the evaluation of the severity of the disorder of two mental patients on the basis of two equally important test scores. The two patients have the same mean score, but whereas one has two moderate scores, the other has one high score (a score indicative of severe disorder) and one low score. If decisions follow a linear compensatory strategy, the evaluations of the two patients would be about the same. If decisions follow a disjunctive strategy, the candidate with the higher scatter receives a higher (more severe) evaluation. If decisions follow a conjunctive strategy, the candidate with the higher scatter receives a lower evaluation.[5]

The Simple Version of the Scatter Model

The simple version of the scatter model represents judgment by two elements: (a) the elevation of each profile, which is a weighted average of the cue values; and (b) the internal scatter of the profile, defined by the variability of the (standardized) cue values around the profile's mean (see Cronbach & Gleser, 1953, for an early treatment of the concepts of elevation and scatter). The influence of the profile's scatter on judgment is indicative of reliance on disjunctive or conjunctive rules. If a disjunctive rule is used, scatter is positively related to judgment; whereas if a conjunctive rule is used, it is negatively related to judgment.

Mathematically, the scatter model is expressed as

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \delta \mathrm{SXT}, \qquad (1)$$

where Y is the judgment, the X_is are the cues, and SXT is the scatter, defined as

---

[3] Note that this description of disjunctive (conjunctive) strategies need not be viewed as a description of the process of the judgment. It can also be viewed as a description of the outcome of the judgment. For example, consider a case in which the stimulus scale is nonlinear, so at low (high) values of the cues, small changes in cue value have a large (small) impact on judgments. In this case, a linear integration process would lead to conjunctive relationships between cues and judgment. Similarly, if at high (low) values of the cues, small changes in cue value have a large (small) impact on judgments, a linear integration process would lead to a disjunctive relationships between cues and judgment (Einhorn, 1970).

[4] These representations of conjunctive and disjunctive rules do not use the concepts of logical *inclusive or* and logical *and,* which are commonly used in defining disjunctive and conjunctive choice strategies. These concepts are somewhat problematic in the context of judgment because they imply that in a conjunctive (disjunctive) strategy, increase (decrease) in cue value above (below) a certain cutoff value does not influence judgment at all. The description of conjunctive and disjunctive strategies allows for changes in cue values to influence judgment across the entire range of possible values, although still retaining the essence of what is meant by conjunction (disjunction) in choice because it suggests that the cues with low (high) values play a major role in the decision. In choice, this is due to the existence of a cutoff value; in judgment, it is due to the dominance of the cue(s) with the lowest (highest) value(s).

For a more formal treatment, consider the two attribute cases. According to the definitions of disjunctive and conjunctive strategy, the relationship between judgment (Y) and cues (X_1 and X_2) could be represented, respectively, as $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \max(X_1, X_2)$ and $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \min(X_1, X_2)$, where Y is the judgment and X_1 and X_2 are the attributes' values. (As a matter of fact, only one equation is necessary to represent both strategies, e.g., the first equation represents a disjunctive strategy if $\beta_3 > 0$ and a conjunctive strategy if $\beta_3 < 0$.) The similarity between the description of disjunctive–conjunctive judgment strategies and the standard definition disjunctive–conjunctive choice strategies becomes apparent by setting $\beta_1 = \beta_2 = 0$.

[5] Note that in the two cue cases the relationship between disjunctive–conjunctive strategies and scatter is evident because the two equations of Footnote 4 can also be written as $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 \mathrm{ABS}(X_1 - X_2)$, where β_3 is positive for a disjunctive strategy and negative for a conjunctive strategy. [See Ganzach & Czaczkes, in press, for a proof. Note that in the two cue cases $\mathrm{ABS}(X_1 - X_2)$ is the scatter and that this expression of scatter is related to the way scatter is expressed in Equations 1, 3, and 5 by a multiplicative constant.]

$$[\sum_{i=1}^{m}(X_i - \bar{X})^2]^{1/2},$$

where $\bar{X}$ is the mean X_i within each profile,

$$\bar{X} = \sum_{i=i}^{n} X_i/n.$$

The X_is are standardized and scaled so that higher values of X_i imply higher Y.

The last term of the equation is a measure of the profile scatter. The value of δ, the scatter coefficient, is indicative of whether conjunctive or disjunctive rules are used. A positive value of δ is indicative of a disjunctive rule, whereas a negative value is indicative of a conjunctive rule. Note that standardization of the cues is necessary so that the value of the scatter does not depend on the scales of the cues. The previous discussion also suggests that to obtain a meaningful scatter term, all cues must be linearly rescaled, if necessary, so their correlations with the judgment would have the same sign.

The simple version of the scatter model was previously used by Brannick and Brannick (1989) to examine conjunction in judgment and by Ganzach (1993, 1994b) to analyze both disjunction and conjunction. It has two advantages over earlier methods of estimating disjunctive and conjunctive strategies (i.e., examination of Einhorn's, 1970, hyperbolic and parabolic models against the linear model). First, it treats judgment strategies as located on a continuum, ranging from conjunctive through linear to disjunctive. This facilitates aggregation of individual strategies for the purpose of examining hypotheses about groups by calculating the average δ (computed for each judge individually) over judges. Second, this model produces a better fit than Einhorn's (1970) models when strategies are "purely" disjunctive or conjunctive (see Brannick & Brannick, 1989; Ganzach & Czaczkes, 1995).

The Multiple-Scatter Version

The simple version of the scatter model does not take into account the structure of the cue information and, in particular, the dimensional structure of this information. No knowledge about the nature of the judgment task is needed to examine for nonlinearity; a simple cookbook approach is used in which scatter information is aggregated and represented by a single, general scatter term.

However, when cues are (perceived to be) organized in dimensions, the nonlinear impact of a cue belonging to a certain dimension may be determined more by its position relative to cues belonging to this dimension than by its position relative to cues belonging to other dimensions. Thus, nonlinearity may be associated more with intradimension scatters than with a general scatter. In addition, nonlinearity may also be associated with interdimension scatter, which represents the scatter between the dimensions. For example, Ganzach (1994a) studied ability judgments based on two dimensions (motivation and intelligence), each characterized by two scores, and showed that the nonlinear intradimension relationships (the gap between the scores) were different from the nonlinear interdimension relationships. These results suggest that the study of nonlinearity may be improved by incorporating information about cue dimensionality.

The version of the scatter model that takes into account the impact of the dimension structure on nonlinearity is labeled the *multiple-scatter version.* This version is expressed as

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \sum_{i=1}^{n} \gamma_j SF_j + \delta SFT, \tag{2}$$

where m is the number of cues, n is the number of dimensions, SF_j is the intradimension scatter of dimension j, and SFT is the interdimension scatter.

SF_j, the intradimension scatter, is defined as

$$SF_j = [\sum_{k=1}^{q_j} (X_k - F_j)^2]^{1/2}, \tag{3}$$

where q_j is the number of cues associated with dimension j (the summation is on the cues associated with dimension j). F_j, the value of dimension j, is derived from an equal weighing average of the attributes associated with the dimension:

$$F_j = \frac{1}{q_j} \sum_{k=1}^{q_j} X_k. \tag{4}$$

SFT, the interdimension scatter, is defined by

$$SFT = [\sum_{j=1}^{n} (F_j - \bar{F})^2]^{1/2}, \tag{5}$$

where $\bar{F}$ is the mean of the dimensions and is defined as

$$\bar{F} = \frac{1}{n} \sum_{j=1}^{n} F_j. \tag{6}$$

In Equations 5 and 6, the F_js are standardized, and if they have a negative correlation with the judgment, they are rescaled to have a positive correlation.

The Deviation Versions: Representing Scatter by Individual Cue Deviations

In the deviation versions, scatter information is represented by individual cue deviations rather than by the average of these deviations. This representation allows for each deviation to have its own weight, instead of assigning the same weight for all deviations.

In the first of the two deviation versions analyzed in this article (labeled the *simple deviation version*), the deviations of the cues are from the mean cue value:

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \sum_{i=1}^{m} \gamma_i D_i, \tag{7}$$

where D_i = ABS($X_i - \bar{X}$). (Again, cues are scaled to have a positive correlation with judgment and are standardized.)

This deviation version is associated with the simple version of the scatter model. Whereas in the simple version of the scatter model the deviations are averaged, represented as SXT, and share the same weight, in this deviation version each deviation is allowed to have its own weight. This may be consequential when the deviation of one cue (e.g., an important cue) has a stronger effect on judgment than the deviation of another cue (e.g., an unimportant cue).

The second deviation version is associated with the multiple-scatter version. In this deviation version, the deviations of the individual cues from the mean of the dimension replace the intradimension scatters, and the deviations of the dimensions from the mean of the dimensions replace the interdimension scatter. This deviation version is expressed as

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \sum_{i=1}^{m} \gamma_i D'_i + \sum_{j=1}^{n} \delta_j DF_j, \quad (8)$$

where $D'_i = ABS(X_i - F_j)$ and $DF_j = ABS(F_j - F)$. Thus, in this version, both the deviations of various cues within each dimension and the deviations of the dimensions from the mean of the dimensions are allowed to have different weights.

Results

The Simple Version of the Scatter Model

The scatter model of Equation 1 was estimated for each of the 29 judges. Prior to the analysis, the two necessary transformations previously described were performed. First, for each judge the correlations between cues and judgment were computed, and the cues whose correlation with the judgment was negative were rescaled by multiplying them by −1. Second, the cues were standardized. Furthermore, for the sake of comparability between various analyses, not only the cues but also the scatter term of each judge were standardized across profiles.

Columns 2 and 6 of Table 1 present the cross-validated multiple correlation (CVMC) of the linear model and of the simple version of the scatter model, respectively. These correlations were obtained by splitting the 861 profiles into two sets through an odd–even split, calculating the CVMC of each of the sets and averaging the two correlations. For 24 out of the 29 judges, the CVMC of the scatter model exceeded the CVMC of the linear model. The hypothesis that the means of the CVMCs of the two models are equal is rejected at $p = .0001$, $t = 5.2$.[6] Thus, it is clear from these data, that the scatter model gives a better fit to the data than the linear model.

Column 10 of Table 1 presents the value of the scatter coefficient (the coefficient of SXT) of each of the judges. It is clear from the table, that for all judges this value is positive. Furthermore, for 25 out of the 29 judges, this value is significantly different than zero ($p < .0001$). These data suggest a consistent use of nonlinear strategy in these clinical judgments—a disjunctive judgment strategy. That is, the data suggest that clinicians assign more weight to the more pathological cues. Thus, not only does the scatter model reveal that judgments are nonlinear, but it also reveals the type of nonlinearity underlying these judgments.

Some alternative linear models. Because a large number of attributes may lead to a degradation in the cross-validated fit of a linear model, I also examined the cross-validated fit of two additional linear models. One was based on the first five MMPI scales entering into each of the stepwise regressions of the judges, and the other was based on the first three scales in such stepwise regressions. The CVMC of these two models are presented, respectively, in columns 3 and 4 of Table 1. These data indicate that the CVMC of the two models was lower than not only the CVMC of the scatter model but also the CVMC of the (11 scales) linear model. The null hypothesis that the CVMC of this latter model is equal to the CVMC of the five- and three-scales linear models was rejected ($t = 3.3$, $p < .005$ and $t = 7.6$, $p < .0001$, respectively).

Normalization. Some concern was raised in the literature regarding the impact of the skewness of the cue distributions on the validity of multiple regression analysis of judgment (Goldberg, 1976). Because the cue distributions in Meehl's data are somewhat skewed, I also examined the impact of scatter on judgment after rescaling the cues by a normalizing transformation (Blom, 1958).

To compare the impact of the transformation, I calculated for each of the judges the partial correlation between the scatter and the judgment (controlling for the linear variance associated with the 11 scales) both before the normalization and after the normalization. The two partial correlations appear, respectively, in columns 15 and 16 of Table 1. It is clear from the table that the partial correlation is lower before the normalization than after the normalization. The difference between the two is significant when $p = .0001$, $t = 15.0$. Thus, normalizing the cues only makes the effect of scatter on judgment more pronounced.[7]

Why does the effect of scatter on judgment increase after normalization? Table 2 presents the correlations between the typical scatter term (in which only two of the MMPI scales, lie [*L*] and hysteria [*Hy*], were multiplied by −1)[8] and the 11 cues. It is clear from the table that the multicollinearity between cues and scatter is high before the normalization and low after it. Thus, the prominence of the scatter effect after normalization is due to decreased multicollinearity.

The Multiple-Scatter Version

The stimuli in Meehl's data are characterized by three clear dimensions. These dimensions are apparent in Table 3, which presents the results of a principle component analysis with a varimax rotation on the 11 scales of Meehl's data. One dimension (labeled F_1) is associated with the scales hypochondriasis (*Hs*), depression (*D*), *Hy*, and psychasthenia (*Pt*). Another di-

[6] In this test, as in the other significant tests involving correlations, the individual correlations were transformed prior to the analyses by Fisher's *r*-to-*z* transformation. In addition, all *t* values reported have 28 degrees of freedom.

[7] Another way by which the effect of normalization can be demonstrated is by comparing the difference in model fit between the scatter model and the linear model before and after normalization. After normalization, the difference between the two is .014 (the CVMC of the scatter model is .762, and the CVMC of the linear model is .748). Before the normalization, the difference between the two is .09 (see Table 1). The null hypothesis of equality between these differences is rejected, $t = 10.1$, $p < .0001$.

[8] Because the computation of the scatter term varies among judges, depending on each judge's cue–judgment correlations, in the typical scatter term only the cues whose correlation with the judgment was negative for the majority of the judges was multiplied by −1. Note also that the typical scatter term is equal to the scatter term of the composite judge, a judge constructed by averaging the judgments of the 29 judges.

Table 1
Comparison of the Alternative Models

	Model fit: Linear models				Model fit: Scatter versions				Regression slopes					Partial correlations	
Judge	11 att.	5 att.	3 att.	3 dim.	Scatter	Multiple scatter	Deviation 1	Deviation 2	SXT	SF_1	SF_2	SF_3	SFT	Before normalization	After Normalization
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	.687	.665	.584	.660	.699	.738	.741	.734	.157	−.143	.254	−.039	.043	.206	.232
2	.819	.807	.779	.789	.819	.835	.835	.835	.081	−.087	.158	.012	.042	.137	.192
3	.707	.698	.686	.657	.728	.736	.723	.731	.192	−.081	.203	.025	.119	.257	.304
4	.703	.684	.649	.681	.724	.736	.731	.730	.192	−.166	.198	.007	.026	.259	.300
5	.756	.751	.725	.716	.761	.786	.791	.790	.125	−.078	.240	−.002	.030	.185	.235
6	.792	.773	.741	.731	.811	.829	.835	.838	.196	−.179	.190	.001	.083	.309	.353
7	.725	.691	.661	.704	.755	.778	.780	.788	.215	−.040	.303	.034	.094	.299	.339
8	.795	.797	.787	.728	.816	.807	.800	.812	.193	.098	.146	.012	.044	.303	.364
9	.806	.799	.766	.771	.828	.827	.820	.837	.200	.027	.151	−.003	.125	.327	.374
10	.712	.717	.707	.694	.726	.711	.715	.730	.146	.047	.109	−.008	.032	.200	.205
11	.836	.831	.822	.792	.841	.840	.835	.841	.115	.000	.120	.002	.054	.203	.270
12	.750	.722	.684	.674	.763	.779	.770	.774	.182	.001	.228	−.001	.093	.262	.317
13	.717	.713	.679	.689	.724	.731	.731	.731	.110	−.062	.149	.008	.066	.154	.219
14	.550	.568	.566	.426	.537	.558	.577	.571	.058	−.136	.085	.014	.082	.068	.106
15	.820	.813	.771	.749	.824	.836	.836	.836	.087	−.104	.147	.000	.062	.146	.236
16	.830	.821	.797	.776	.836	.861	.860	.859	.113	−.088	.222	−.028	.043	.194	.244
17	.729	.723	.623	.695	.728	.737	.742	.741	.056	−.096	.120	.036	.003	.079	.161
18	.842	.843	.819	.752	.854	.861	.857	.863	.134	.008	.193	.002	.050	.240	.301
19	.772	.771	.756	.732	.786	.800	.795	.811	.159	−.060	.189	−.020	.092	.245	.273
20	.760	.751	.729	.747	.768	.766	.758	.759	.131	.010	.165	−.005	.088	.187	.244
21	.836	.834	.825	.800	.845	.836	.835	.839	.143	.046	.104	.013	.061	.255	.305
22	.785	.782	.762	.776	.800	.806	.807	.818	.165	.037	.190	−.019	.056	.261	.300
23	.729	.730	.712	.624	.728	.742	.746	.744	.096	−.100	.145	−.001	.029	.136	.200
24	.682	.678	.660	.633	.692	.700	.687	.712	.129	−.074	.138	−.033	.126	.168	.206
25	.829	.831	.819	.773	.832	.831	.839	.842	.106	.047	.094	−.021	.031	.184	.225
26	.759	.776	.742	.696	.760	.757	.764	.766	.094	−.016	.112	.014	.048	.143	.210
27	.719	.721	.713	.701	.723	.724	.725	.723	.096	−.022	.108	.043	.029	.135	.165
28	.804	.792	.726	.766	.803	.824	.821	.824	.058	−.097	.183	.005	.004	.094	.175
29	.767	.757	.730	.727	.767	.781	.784	.785	.061	−.088	.150	.002	.031	.092	.137
Mean	.759	.753	.725	.712	.768	.778	.777	.782	.131	−.048	.165	.002	.058	.197	.248

Note. att. = attribute; dim. = dimension; SXT = overall scatter (the scatter of the simple version); SF_1 = intradimension scatter of the first dimension; SF_2 = intradimension scatter of the second dimension; SF_3 = intradimension scatter of the third dimension; SFT = interdimension scatter.

Table 2
Correlation Between the Typical Scatter and MMPI Scales

Scale	Before normalization	After normalization
L	.03	−.03
F	.25***	.09*
K	−.04	−.06
Hs	.14***	−.04
D	.09*	−.03
Hy	.08*	−.06
Pd	.11**	−.02
Pa	.18***	−.02
Pt	.11**	−.05
Sc	.21***	.00
Ma	.11**	.00

Note. MMPI = Minnesota Multiphasic Personality Inventory; *L* = lie; *F* = eccentricity; *K* = defensiveness; *Hs* = hypochondriasis; *D* = depression; *Hy* = hysteria; *Pd* = psychopathic deviate; *Pa* = paranoia; *Pt* = psychasthenia; *Sc* = schizophrenia; *Ma* = hypomania.
* $p < .01$. ** $p < .001$. *** $p < .0001$.

mension (F_2) is associated with the scales eccentricity (*F*), psychopathic deviate (*Pd*), paranoia (*Pa*), schizophrenia (*Sc*), and hypomania (*Ma*). The third dimension (F_3) is associated with the scales *L* and defensiveness (*K*). (The eigenvalues of the defensiveness three factors are 3.2, 2.7, and 1.6, respectively.)

To estimate the parameters of the multiple-scatter version, the value of each of the three dimensions was determined by averaging over the scales associated with it (Equation 4; actual dimensions were used as a proxy for perceived dimensions), the value of each intradimension scatter was determined from the deviations of the dimension cues from their average (Equation 3), and the value of the interdimension scatter was determined from the deviations of the dimensions' values from their mean (Equation 5). In the calculation of this latter scatter, the third dimension was rescaled for 28 of the judges by multiplying it by −1 because, except for one judge (Judge 20), the correlation between F_3 and the judgment was negative.

Following these calculations, the multiple-scatter version of Equation 2 was estimated for each judge. For the sake of comparability, the scatter terms were standardized across profiles. Column 7 of Table 1 presents the CVMC of the multiple-scatter model. It is clear from the results that the fit of the multiple-scatter version exceeds the fit of the simple version. For 21 of the judges, the CVMC of the multiple-scatter version was higher than the CVMC of the simple version. The difference between the two was significant when $p = .0005$, $t = 4.1$.

Columns 11, 12, 13, and 14 of Table 1 present the coefficients of SF_1, SF_2, SF_3, and SFT. The data indicate that the mean coefficients of SF_2 and SFT are positive ($t = 17.1$, $p < .0001$ and $t = 9.4$, $p < .0001$, respectively). Thus, in regard to the interdimension scatter and to SF_2, the judgments are disjunctive. However, the mean coefficient of SF_1 is significantly negative, $t = 3.7$, $p < .001$. Thus, in regard to SF_1, the judgments are conjunctive. This difference between the nonlinear relationships associated with the scales of F_1 and the nonlinear relationships associated with the scales of F_2 is an important reason for the superior fit of the multiple-scatter version over the simple version. The simple version does not take into account the dimension specificity of the nonlinear relationships between attributes and judgment and "aggregate" over both disjunctive and conjunctive relationships.[9] As a result, the nonlinear variance extracted by this version is rather low. However, in the multiple-scatter version, there is no aggregation over conflicting nonlinear relationships. In this version, the nonlinear (conjunctive) variance associated with the scales of F_1 and the nonlinear (disjunctive) variance associated with the scales of F_2 are extracted separately.

Finally, because the data contain three distinct dimensions, an additional linear model, which is worth examining, is a linear model that includes only the three dimensions:

$$Y = \alpha + \sum_{j=1}^{n} \beta_j F_j. \quad (9)$$

The potential advantage of this model is that, on the one hand, it is associated with little degradation on cross validation, but on the other hand it incorporates more information than the three- and five-attribute linear models described earlier. However, the results of the cross-validation indicated that the CVMC of this three-dimension linear model was lower, not only in comparison to the multiple-scatter model but also in comparison to the 11-scale linear model, $t = 9.9$, $p < .0001$. (Column 5 in Table 1 presents this model's CVMC.)

The Deviation Versions

The CVMC of the simple deviation version (associated with the simple version of the scatter model) is presented in column 8 of Table 1, and the CVMC of the deviation version associated with the multiple-scatter version is presented in column 9 of the table. It is apparent from the data that the two deviation versions give a better fit to the data than the linear model.

To test whether the representation of scatter by individual deviations, rather than by the average of these deviations, improves model fit, I compared each of the two deviation versions with its control version. In both cases, the null hypothesis was rejected. The difference between the simple version of the scatter model and the simple deviation version was significant when $p = .0001$ ($t = 6.1$). The difference between the multiple-scatter version and the deviation version associated with it was significant when $p = .003$ ($t = 3.3$).

Discussion

The analyses presented in this article indicate that the various versions of the scatter model can detect nonlinearity in clinical

[9] This aggregation results in an overall pattern of conjunctive relationships (the scatter coefficient in this version is negative) because the interdimension relationships are disjunctive and because the disjunctive relationships associated with the scales of F_2 are stronger than the conjunctive relationships associated with the scales of F_1 (the incremental variance associated with the former dimension is smaller than the incremental variance associated with the latter dimension).

Table 3
Factor Structure of the MMPI Scales

Scale	Factor 1	Factor 2	Factor 3
L	.08	.03	.73
F	.18	.73	−.28
K	−.08	−.15	.85
Hs	.81	.06	.17
D	.85	.10	−.25
Hy	.85	.08	.22
Pd	.20	.70	.15
Pa	.45	.61	−.11
Pt	.76	.40	−.20
Sc	.55	.71	−.06
Ma	−.20	.71	−.04

Note. Entries are factor loadings. MMPI = Minnesota Multiphasic Personality Inventory; *L* = lie; *F* = eccentricity; *K* = defensiveness; *Hs* = hypochondriasis; *D* = depression; *Hy* = hysteria; *Pd* = psychopathic deviate; *Pa* = paranoia; *Pt* = psychasthenia; *Sc* = schizophrenia; *Ma* = hypomania.

judgments that are based on representative stimuli. These results are in contrast to the failures of previous models, such as the hyperbolic and the parabolic models (Einhorn, 1970) or the sign and the quadratic models (Wiggins & Hoffman, 1968), to find systematic reliance on nonlinear strategies in the same data.

In assessing the meaning of these results, remember that the incremental fits over the linear model, although highly significant, are quite small. Thus, for practical reasons, such as replacing linear models by scatter models in "bootstrapping" expert judgment, the implications of these findings are limited.[10] However, these findings are quite relevant to the understanding of the cognitive processes underlying judgment because (as a result of the robustness of linear models) even small changes in fit may reflect important changes in underlying processes. Thus, the results of the article not only reveal that MMPI-based clinical judgments are nonlinear but also reveal systematic patterns of nonlinearity in regard to the interdimension information and in regard to two categories of intradimension information. The results also reveal that the analysis of nonlinearity by a model in which scatter information is aggregated and represented by a single, general scatter term (the analysis on the basis of the simple version of the scatter model) may obscure fine-grained nonlinear relationships that are revealed only in more detailed analyses (e.g., the analysis on the basis of the multiple-scatter version).

Why did the current attempt to find a systematic pattern of nonlinearity in Meehl's data succeed where previous work have failed? There are two reasons for the success of the scatter model in comparison with other nonlinear models: a statistical reason and a psychological reason.

From a statistical point of view, the scatter model is more desirable than Einhorn's (1970) models, as well as the models offered by Wiggins and Hoffman (1968), because its important results (e.g., model fit) do not change under linear transformation of the scales. Einhorn's models, and Wiggins and Hoffman's models, lack this property and, therefore, lead to arbitrary findings (Goldberg, 1971).

From a psychological point of view, I believe that the success of the various forms of the scatter model and, in particular, the success of the multiple-scatter version is associated with the fact that they capture important features of intuitive nonlinear strategies. First, to a large extent, nonlinear strategies are associated with intuitive theories concerning both inter- and intradimension integration. Second, the conjunctive–disjunctive rules are fundamental, or prototypical, nonlinear rules that represent many features of nonlinear information processing. Because the scatter model supplies a good mathematical representation of these rules, it can successfully model task-specific and even person-specific nonlinear strategies.

To demonstrate the prototypicality of disjunctive and conjunctive strategies, consider some of the nonlinear signs routinely used in MMPI judgments, which are described by Goldberg (1965). Two of these signs are labeled the *high point rules* (Signs 61 and 62 in Goldberg's list). One of them suggests that if a patient's highest score is on a scale originally derived from neurotic patients (*Hs, D, Hy,* or *Pt*), the patient should be classified as neurotic; if the high point falls on any other scale, he or she should be classified as psychotic. The other sign is a similar classification system based on the profile's low score. These signs are clearly examples of disjunctive and conjunctive rules. (Note, however, that the judgments in Meehl's data are consistent only with the first of these two rules). Other examples of signs that represent conjunctive–disjunctive rules are those that correspond to the relative elevation of the dimensions (Signs 25–30) and to the relative elevation within the dimension (Signs 31–33).

The fundamental nature of the disjunctive–conjunctive rules discussed in the article can also be demonstrated quantitatively by comparing the predictive power of the commonly used nonlinear signs described by Goldberg (1965) against the predictive power of the disjunctive–conjunctive signs used in this article (the various forms of scatter). Goldberg correlated the MMPI signs described in his earlier article (Goldberg, 1965) with the residuals obtained from regressing the average judgments of the 29 clinicians on the 11 cues. He found a relatively small number of significant correlations, the largest of them equaling .23.[11] I repeated Goldberg's (1971) procedure and correlated the residuals with the basic five disjunctive–conjunctive signs analyzed in the current article (i.e., SXT, SFT, SF_1, SF_2, and SF_3[12]). Out of these five correlations, four were significant. Two of them, the correlation between the residuals and SF_2 and the correlation

[10] Keep in mind that Meehl's data do not include replicates, so it is impossible to estimate how much of the gap between the systematic variance and the linear variance is explained by the nonlinear terms. However, in this respect, the nonlinear variance is most likely not negligible because (due to the robustness of the linear model) the systematic variance and the linear variance are most likely quite close to each other (see Brehmer & Brehmer, 1988).

[11] I was not able to determine the number of correlations Goldberg (1971) considered. Thus, for example, from the 13 significant correlations, only 6 appear in the table that describe the basic diagnostic signs (Goldberg, 1965, p. 4). It is clear, however, that many such correlations were considered.

[12] Because SXT and SFT are defined individually for each judge, the typical values of these scatters (see Footnote 8) were used in this calculation.

between the residuals and SXT, exceeded the highest correlation found by Goldberg (1971). The former correlation was .39 and the latter .30.

One important feature of the various versions of the scatter model is the representation of nonlinearity as a continuum ranging from disjunctive to conjunctive. As a result, the model allows for aggregating over individuals to arrive at conclusions concerning the various types of nonlinearity in the population. This advantage of the scatter model is demonstrated in the analyses based on mean scatter coefficients. These analyses not only show the existence of nonlinearity but also uncover distinct psychological processes underlying this nonlinearity and, in particular, indicate that these processes are dimension specific. The integration of the cues associated with one dimension follows a disjunctive rule, whereas the integration of the cues associated with the other dimension follows a conjunctive rule.

The psychological interpretation of the nonlinear rules associated with these two dimensions is especially interesting because these dimensions have distinct characteristics: One dimension (F_1) is associated with the neurotic scales of the MMPI and the other (F_2) is associated with the psychotic scales. Therefore, although the two nonlinear rules are statistically different, they may be psychologically similar in that both are disjunctive vis-à-vis the dimension. The high-value neurotic scales receive relatively higher weight in judging the degree of neurosis, whereas the high-value psychotic scales receive relatively high values in judging the degree of psychosis.

References

Ben-David, A., & Mandel, J. (in press). Classification accuracy: Machine learning versus explicit knowledge acquisition. *Machine Learning.*

Blom, G. (1958). *Statistical estimates and transformed beta variables.* New York: Wiley.

Brannick, M. T., & Brannick, J. P. (1989). Nonlinear and noncompensatory processes in performance evaluation. *Organizational Behavior and Human Decision Processes, 44,* 97–122.

Brehmer, B., & Brehmer, A. (1988). What have we learned about human judgment from thirty years of policy capturing? In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The social judgment theory view.* Elsevier Science: Amsterdam.

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

Dawes, R. M. (1964). Social selection bases on multidimensional criteria. *Journal of Abnormal and Social Psychology, 68,* 104–109.

Dawes, R. M., & 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. (1971). The use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Performance, 6,* 1–27.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Performance, 7,* 86–106.

Einhorn, H. J., Komorita, S. S., & Rosen, B. (1972). Multidimensional models for the evaluation of political candidates. *Journal of Experimental Social Psychology, 8,* 58–73.

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

Ganzach, Y. (1994a). Inconsistency and uncertainty in multi-attribute judgment of human performance. *Journal of Behavioral Decision Making, 7,* 193–211.

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

Ganzach, Y., & Czaczkes, B. (1995). On detecting nonlinear noncompensatory judgment strategies: Comparison of alternative regression models. *Organizational Behavior and Human Decision Processes, 61,* 168–177.

Ganzach, Y., & Czaczkes, B. (in press). The learning of natural configural strategies. *Organization Behavior and Human Decision Processes.*

Goldberg, L. R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis versus neurosis from the MMPI. *Psychological Monographs, 79*(9, Whole No. 602).

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. (1976). Man versus model of man: Just how conflicting is that evidence. *Organizational Behavior and Human Performance, 16,* 13–22.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin, 57,* 116–131.

Johnson, E. J. (1988). Expertise and decision under uncertainty: Performance and process. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 209–228). Hillsdale, NJ: Erlbaum.

Johnson, P. E., Hassebrock, F., Duran, A. S., & Moller, J. (1982). Multimethod study of clinical judgment. *Organizational Behavior and Human Performance, 30,* 201–230.

Kleinmuntz, B. (1963). MMPI decision rules for the identification of college maladjustment: A digital computer approach. *Psychological Monograph, 77*(14, Whole No. 577).

Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota Press.

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

Ogilvie, J. R., & Schmitt, N. (1979). Situational influences of linear and nonlinear use of information. *Organizational Behavior and Human Performance, 23,* 292–306.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance, 6,* 649–744.

Wiggins, N., & Hoffman, P. J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology, 73,* 70–77.

Wright, P. (1974). The harassed decisionmaker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology, 59,* 555–561.

Received December 15, 1993
Revision received April 25, 1995
Accepted April 25, 1995 ■