Research Article

NONLINEAR MODELS OF CLINICAL JUDGMENT: Communal Nonlinearity and Nonlinear Accuracy

Yoav Ganzach

Tel Aviv University, Tel Aviv, Israel

Abstract—*Despite our intuition that representative expert judgments are highly nonlinear, previous studies have shown only little, if any, nonlinearity in such judgments. The current study presents a method for assessing nonlinearity in judgment that is based on estimating communal nonlinearity—the systematic nonlinearity shared by the community of judges. The article also examines the predictive accuracy of communal nonlinearity, and compares it with the corresponding linear accuracy.*

A considerable portion of the decision-making literature has been devoted to empirical research examining nonlinearity in multiattribute judgment. The results of this research have been mixed. On the one hand, studies based on nonrepresentative, orthogonal stimuli, conducted within Anderson's (1981) information-integration paradigm, and using primarily nonexpert participants have documented strong nonlinearity in judgment (e.g., Birnbaum, Coffey, Mellers, & Weiss, 1992; Meyer, 1987; Weber, 1994). On the other hand, studies based on nonorthogonal, representative stimuli, conducted within the regression paradigm (Slovic & Lichtenstein, 1971), and using primarily expert participants have documented little, if any, nonlinearity in judgment (e.g., Billing & Marcus, 1983; Brannick & Brannick, 1989; Goldberg, 1971; Wiggins & Hoffman, 1968; see Brehmer & Brehmer, 1988, for a review). These latter findings are surprising because both process-tracing studies (e.g., E.J. Johnson, 1988; P.E. Johnson, Hassebrock, Duran, & Moller, 1982) and people's intuition about their judgments (e.g., Meehl, 1954) suggest that judgments are highly nonlinear.

More recently, however, I showed (Ganzach, 1995) that some versions of a nonlinear regression model labeled the scatter model, which takes into account the internal variation of attribute values (i.e., the degree of inconsistency between the profile's attributes), do succeed in explaining more variance in judgment than a simple linear model (see also Brannick & Brannick, 1989; Ganzach & Czaczkes, 1995). Nevertheless, the additional variance explained by the scatter model was not large. The best version of this model provided an incremental variance (ΔR^2) of about .03 over the linear model, which amounted to an increase of only 5% of the explained variance. Thus, although this work demonstrates a significant nonlinear element in judgment, this element is rather small.

Why was the additional fit of regression models that attempted to take into account nonlinearity in judgment so small and hard to detect? There are three possible answers to this question. One is that there is indeed little nonlinearity in human judgment—that despite our intuition that judgments are often highly configural, in reality the processes underlying judgment are linear. Another explanation is the robustness of the linear model—the fact that linear models provide good fit to the data even if the "true" model is highly nonlinear (Dawes & Corrigan, 1974). Finally, a third explanation is that so far, all the models that have attempted to examine nonlinearity in judgment have required the specification of the nonlinear process underlying judgment; that is, the researcher has had to represent the nonlinear processes as a function of the attribute values. If this specification was inaccurate, the nonlinear variance could not be large.

The current study takes a new approach to modeling nonlinearity in judgment. This approach relies on the residuals from a linear model of the judgment of a reference group to specify nonlinearity in the judgment of other judges. If the nonlinear strategies of these judges are similar to the nonlinear strategies of the reference group, a representation of the nonlinear processes in their judgment could be obtained without specifying these processes.

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

NONLINEAR COMMUNALITY

In judgment modeling, a judgment of a multiattribute profile is often expressed as the sum of a linear combination of the attributes' values and a residual. The residual can be divided into two major parts, one representing nonlinear rules underlying the judgment and the other representing error. The nonlinear part can further be divided into two portions, one representing *communal nonlinearity*—systematic nonlinearity shared by the community of judges—and the other representing *idiosyncratic nonlinearity*—nonlinearity specific to the individual judge. Note that as defined here communal nonlinearity is profile-specific, but not judge-specific, whereas idiosyncratic nonlinearity is both profile- and judge-specific. A formal definition of these concepts appears in the appendix.

It is clear how the attribute values of each profile could be used in modeling the judgments of each of the judges. However, it is less clear how communal nonlinearity could be used as well. Therefore, the assessment of this communal nonlinearity, and its use in judgment modeling, is at the center of the modeling approach suggested in this article.

In principle, the communal nonlinearity of a profile can be assessed by averaging the judgment residuals (from a linear model of the judgment) of a group of judges. This averaging "cancels out" the error component as well as the idiosyncratic nonlinearity component, and provides a good estimate of the nonlinearity that is common to the population of judges. However, to use communal nonlinearity in modeling multiattribute judgments, it is necessary to know, a priori, its value for each of the multiattribute profiles. Therefore, in modeling the judgment of judge *j*, the appropriate estimate for the communal nonlinearity of a profile is obtained by averaging the judgment residuals (from the linear model of the attributes) of this profile across all other judges. That is, although in principle underlying communal nonlinearity is not judge-specific, and is only profile-specific, for each judge I compute a slightly different communal nonlinearity that can be labeled the "average residual of the other judges," and this value is not only profile-specific, but also judge-specific. A formal definition of this variable appears in the appendix.

In sum, although in principle communal nonlinearity is a function of the attribute values (it may involve curvilinear, interactive, or any other nonlinear use of attributes), my measure of communal nonlinearity allows for its estimation without specifying this function. As a result, this measure can provide a good method for estimating nonlinearity in judgment by circumventing the need to specify the functional form of this nonlinearity.

NONLINEAR COMMUNALITY AND NONLINEAR ACCURACY

Communal nonlinearity is associated not only with the average judgment residual of a group of judges, but also with the residual of the (linear) model of the *composite judge*—a "judge" whose judgments are the average of the judgments of all judges. The average judgment residual of a group of judges and the residual of the composite judge are exactly equal.

In this section, I discuss the predictive accuracy of judgments—the extent to which a judgment predicts the criterion. Because it is more convenient to analyze the relationship between communal nonlinearity and the predictive accuracy of judgment in terms of the residual of the composite judge than in terms of the average judgment residual of the judges, I develop the discussion of nonlinear accuracy in terms of the former rather than the latter.

A number of studies have shown that nonlinear elements of judgment may have predictive accuracy (Einhorn, 1974; Ganzach, 1998; Goldberg, 1970). That is, they have shown that the correlation between the criterion and the residual of the linear model of a judgment is significantly different from zero. However, all these studies concentrated on the nonlinear accuracy of the individual judges, and none examined whether aggregating over judges improves nonlinear accuracy. That is, none of them examined the nonlinear accuracy of the composite judge (i.e., the nonlinear accuracy of communal nonlinearity). The preceding analysis suggests that if the nonlinear rules shared by the judges have ecological validity, communal nonlinearity is the part of the residual that underlies nonlinear accuracy, because it is communal nonlinearity that captures the systematic nonlinear rules used by the judges. In other words, this line of reasoning suggests that if the source of individual nonlinear accuracy is the nonlinear rules an individual judge shares with the community of judges (rather than his or her idiosyncratic nonlinearity), then the nonlinear accuracy of the composite judge will be higher than the nonlinear accuracy of most of the individual judges.

Within this context, it is interesting to compare the composite judge with the individual judges in terms of two types of predictive accuracy: (a) linear accuracy, the predictive accuracy of the modeled part of the judgment, the part that can be modeled by a linear model; and (b) residual accuracy, the predictive accuracy of the residual part of the judgment. (Formal definitions of linear accuracy and residual accuracy appear in the appendix.) This comparison is interesting because in his seminal work, Goldberg (1970) showed that the linear accuracy of the composite judge is about the average linear accuracy of the individual judges. However, the preceding discussion suggests that the residual accuracy of the composite judge is superior to the residual accuracy of the individual judges.

METHOD

Much of the research aimed at studying nonlinearity in judgment has been based on data collected by Meehl in the 1950s (Meehl, 1959). These data included judgments of 861 patients on the basis of their profiles (i.e., scores on eight clinical scales and three validity scales) on the Minnesota Multiphasic Personality Inventory (MMPI). The judgments were made on an 11-step forced normal distribution ranging from *least psychotic* (1) to *most psychotic* (11). They were obtained from 13 clinical psychologists and 16 clinical psychology trainees.

The data also included the criterion for each patient—the diagnosis given to the patient in the clinic in which he or she received treatment. Forty-seven percent of the patients were diagnosed as psychotics, and 53% were diagnosed as neurotics. Each patient's diagnosis was based primarily on information about the patient's past and present behavior, which was collected by the clinic's staff.

RESULTS

Communal Nonlinearity as a Predictor of Individuals' Judgments

For each judge (e.g., Judge 1) and each profile, I calculated the average residual of the other judges (e.g., the average residual of Judges 2–29). Subsequently, for each judge, I used this average residual as an additional variable, along with each of the 11 MMPI scales, to predict the responses of that judge. Column 2 of Table 1 presents, for each of the 29 judges, the cross-validated multiple correlation (CVMC) of this model, whereas column 3 presents the CVMC of the linear model. These correlations were calculated by splitting the 861 profiles into two sets through an odd-even split, calculating the CMVC of each of the sets, and averaging the two correlations.[1]

It is clear that the CVMC of the model that includes communal nonlinearity far exceeds the CVMC of the linear model. Whereas the average CVMC of the linear model is .76 (SD = 0.06), the average CVMC of the model that includes communal nonlinearity is .84 (SD = 0.06). The difference is highly significant, $t(28) = 16.6$, $p < .0001$. It is clear from these data that the judgment variance explained by communal nonlinearity is not at all trivial; it amounts to 22% of the linear variance.

Column 4 in Table 1 presents the CVMC of the multiple scatter model—the "best" nonlinear model that was applied to Meehl's (1959) data (as well as to other sets of data; see Ganzach & Czaczkes, 1995). The data in this column suggest that the nonlinear variance explained by the scatter model is small compared with the nonlinear variance explained by communal nonlinearity.

On the Sources of Communal Nonlinearity

Goldberg (1971) suggested that the nonlinearity in Meehl's (1959) data arises from specific configural rules that were popular in MMPI-based clinical judgments at the time Meehl's experiment was conducted, and in particular the high-point rules, the number of clinical scales on which the patient scores below 45, and the variance of the Paranoia, Psychasthenia, and Schizophrenia scales (see Goldberg, 1965, for a quantitative specification of these rules). However, an al-

1. Note that a computationally simpler method for obtaining the CVMC associated with adding communal nonlinearity is to calculate, for each judge (e.g., Judge 1), the mean judgment of the other judges (e.g., Judges 2–29) and use it in addition to the 11 MMPI scales to predict the responses of that judge (i.e., Judge 1).

Table 1. *Model fit and accuracies of individual judges*

	Model fit			Predictive accuracy	
Judge	Linear + communal nonlinearity	Linear	Scatter	Nonlinear	Linear
1	.78	.69	.73	.11	.36
2	.87	.82	.84	.01	.35
3	.81	.71	.73	.08	.34
4	.76	.70	.73	.05	.39
5	.88	.76	.79	.12	.33
6	.86	.79	.84	.10	.34
7	.83	.73	.79	.13	.36
8	.87	.80	.81	.00	.32
9	.87	.81	.84	.06	.24
10	.81	.71	.73	.08	.17
11	.88	.84	.84	.07	.21
12	.85	.75	.77	.07	.33
13	.81	.72	.73	.05	.31
14	.60	.55	.57	.05	.20
15	.89	.82	.84	.08	.41
16	.91	.83	.86	.07	.32
17	.79	.73	.74	.06	.43
18	.90	.84	.86	.09	.39
19	.85	.77	.81	.07	.30
20	.85	.76	.76	.08	.24
21	.88	.84	.84	.01	.16
22	.87	.79	.82	.09	.25
23	.79	.73	.74	.05	.36
24	.79	.68	.71	.06	.38
25	.88	.83	.84	.02	.24
26	.83	.76	.77	.05	.37
27	.78	.72	.72	.08	.22
28	.89	.80	.82	.09	.38
29	.85	.77	.86	.08	.32
Average	.84	.76	.78	.07	.31
Composite	—	.89	.92	.12	.33

Note. Correlations higher than .07 are significant, $p < .05$. The model fit of linear + communal nonlinearity cannot be calculated for the composite judge because it is estimated by the average residual of the other judges, which is a judge-specific parameter.

ternative explanation is that nonlinearity in judgment often arises from a general configural confirmatory strategy, and in particular a value-dependent weighing strategy (Ganzach, 2000, Fig. 1). Within the context of Meehl's experiment, this strategy implies that in integrating the neurotic (psychotic) scales to arrive at an overall evaluation of the neurotic (psychotic) tendency of the patient, the judges gave the most neurotic (psychotic) information relatively high weight (see Ganzach, 1998, for a quantitative specification of this strategy).

To examine the nature of the nonlinearity detected by my model, I regressed communal nonlinearity on the main sources of nonlinearity suggested by these two alternative explanations. The results of this regression are given in Table 2. It is clear from these results that the nonlinearity in Meehl's (1959) data is associated both with specific configural rules and with general confirmatory configural strategies, and in particular with the configural integration of the psychotic information.

Finally, note that the overall fit of the model presented in Table 2, although highly significant ($p < .001$), was not large ($R^2 = .19$). Although Goldberg (1971) and I (Ganzach, 2000) suggested some additional, less important sources for nonlinearity, the omission of these sources from the model does not explain its moderate fit, as the additional fit of these sources is small. Rather, in my view, the reason for the moderate fit is the inherent difficulty of identifying and quantifying the numerous complex nonlinear rules and strategies that are often used in clinical judgment.

Nonlinear Communality and Nonlinear Accuracy

Column 5 of Table 1 presents the correlation between the residual part of the judgment and the criterion. These correlations are nearly all lower than .12, the corresponding correlation of the composite judge, $t(28) = 9.1$, $p < .0001$, for the null hypothesis that the correlation of the composite judge equals the average correlation of the individual judges. In fact, the correlation of the composite judge is higher than the correlations of all but one of the individual judges. For compari-

Table 2. *The sources of communal nonlinearity*

Source of nonlinearity	β	Partial *r*
High-Point Rule 1	.11*	.10
High-Point Rule 2	.01	.01
Number of clinical scales on which the patient scores below 45	.23***	.23
Variance of Paranoia, Psychasthenia, and Schizophrenia scales	.07*	.07
Configural integration of neurotic information	−.12**	−.11
Configural integration of psychotic information	.30***	.26

Note. $N = 861$. $R^2 = .19$.
*$p < .05$. **$p < .001$. ***$p < .0001$.

son, column 6 presents the correlation between the predicted part of the judgment and the criterion. In this case, the correlation of the composite judge—.33—is about the average correlation of the individual judges, $t(28) = 0.07$, $p > .9$. It is higher than the correlation of 14 of the judges and lower than the correlation of 13 of the judges.

DISCUSSION

The current study demonstrates larger nonlinearity than has been previously detected in real-world, representative judgments in general, and in Meehl's (1959) data in particular. Thus, it is clear from these findings that representative judgments can be highly nonlinear. The fact that little nonlinearity was found in previous research may be the result of inadequate methods, and not necessarily the result of any linear characteristics of the judgment process.

However, it is important to note that the present method for estimating nonlinearity in judgment does not constitute a model of this nonlinearity in the sense that it does not express nonlinearity in terms of the attributes of the judged object. Therefore, one limitation of the method is that it does not provide information about the nature of the nonlinearity. Another limitation is that it does not allow predictions regarding the judgment of objects not included in the original set of objects used to elicit the judgments. Thus, communal nonlinearity cannot be used in expert systems that are based on regression models of judgments, because such systems rely on the information obtained from the judgments of one set of objects to make predictions regarding the entire universe of objects.

However, examination of communal nonlinearity is worthwhile in developing regression-based expert systems, because it provides the developer with an idea about the adequacy of the systems. If communal nonlinearity explains a nontrivial part of the variance in judgments, systems that are based on linear models are inadequate if the criterion for adequacy is maximization of explained variance. On the other hand, if the judgment variance explained by communal nonlinearity is small, linear systems are adequate.

Finally, examination of communal nonlinearity is worthwhile when trying to understand judgment processes. First, it can provide researchers with an idea about the importance of nonlinear strategies when studying clinical judgment. Second, when some nonlinear rules are known, modeling nonlinear communality as a function of these rules can provide researchers with information regarding the relative importance of these rules in the nonlinear portion of judgments.

Acknowledgments—This research was supported by a grant from the Israel Institute of Business Research. I would like to thank Lewis R. Goldberg, William M. Grove, and Moshe Leshno for many helpful comments.

REFERENCES

Anderson, N.H. (1981). *Foundations of information integration theory.* New York: Academic Press.
Billing, R.S., & Marcus, A.A. (1983). Measures of compensate and noncompensatory behavior: Process tracing versus policy capturing. *Organizational Behavior and Human Decision Processes*, *31*, 331–352.
Birnbaum, M.H., Coffey, G., Mellers, B.A., & Weiss, R. (1992). Utility measurement: Configural weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 331–346.
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. Joyce (Eds.), *Human judgment: The social judgment theory view* (pp. 75–113). Amsterdam: Elsevier Science.
Dawes, R.M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106.
Einhorn, H.J. (1974). Cue definition and residual judgment. *Organizational Behavior and Human Performance*, *12*, 30–49.
Ganzach, Y. (1995). Nonlinear models of clinical judgment: Meehl's data revisited. *Psychological Bulletin*, *118*, 422–429.
Ganzach, Y. (1998). Nonlinear models in decision making: The diagnosis of psychosis versus neurosis from the MMPI. *Organizational Behavior and Human Decision Processes*, *74*, 53–61.
Ganzach, Y. (2000). The weighing of pathological and non-pathological information in clinical judgment. *Acta Psychologica*, *104*, 87–101.
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.
Goldberg, L.R. (1965). Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI. *Psychological Monographs*, *79*(9, Whole No. 602).
Goldberg, L.R. (1970). Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. *Psychological Bulletin*, *73*, 422–432.
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.
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 human judgment. *Organizational Behavior and Human Performance*, *30*, 201–230.
Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota.
Meehl, P. (1959). A comparison of clinicians with five statistical methods of identifying psychotic MMPI profiles. *Journal of Counseling Psychology*, *6*, 102–109.
Meyer, R.J. (1987). The learning of multi-attribute judgment policies. *Journal of Consumer Research*, *14*, 155–173.
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.
Weber, E.U. (1994). From subjective probability to decision weights: The effect of asymmetric loss function on the evaluation of uncertain outcomes and events. *Psychological Bulletin*, *115*, 228–242.
Wiggins, N., & Hoffman, P.J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology*, *73*, 70–77.

(RECEIVED 5/2/00; REVISION ACCEPTED 12/19/00)

APPENDIX

Formal Definitions of Communal and Idiosyncratic Nonlinearity

Consider a situation in which J judges make L ratings regarding each of N multiattribute profiles, each profile consisting of K attributes. Let j, i, and l be the indices of the judge, profile, and replicate rating, respectively, and let k be the index of the attribute. Thus, y_{ijl} is the lth rating of the ith profile of the jth judge, and V_{ik} is the measurement of the kth attribute in the ith profile.

Now consider the following nonlinear regression of y_{ijl} on the K cues:

$$y_{ijl} = \lambda_{ij} + \theta_i + \delta_{ij} + \in_{ijl},$$

where the regression function is given by

$$E_l[y_{ijl} | (V_{ik})] = \lambda_{ij} + \theta_i + \delta_{ij}.$$

$(E_l[\cdot])$ denotes the expectation of the bracketed quantity, taken over the subscript l, λ_{ij} is the linear regression of judgments on attribute values, for the jth judge's ratings of the ith individual, θ_i is the communal (across-judge) nonlinearity, δ_{ij} is the idiosyncratic (judge-specific) nonlinearity, and $\in_{ijl}$ is the random, residual error.

Formally, we define

$$\lambda_{ij} \overset{def}{\equiv} \alpha_j + \sum_{k=1}^{K} \beta_{jk} V_{ik},$$

$$\theta_i \overset{def}{\equiv} E_j E_l[y_{ijl} - \lambda_{ij}],$$

$$\delta_{ij} \overset{def}{\equiv} E[y_{ijl} - \lambda_{ij} - \theta_i],$$

and

$$\in_{ijl} \overset{def}{\equiv} y_{ijl} - E_l[y_{ijl}] = y_{ijl} - \lambda_{ij} - \theta_i - \delta_{ij},$$

where α_j is the linear regression intercept for the jth judge, and β_{ij} is the linear regression slope for the jth judge and the ith subject, for predicting the judgments from the attribute values.

Formal Definition of the Average Residual of the Other Judges

The estimate of communal nonlinearity, labeled "average residual of the other judges," is slightly different for each judge. It is given by

$$\hat{\theta}_{ij'} = M_{j \# j'} M_l [y_{ijl} - \hat{\lambda}_{ij}],$$

where $M[\cdot]$ is the mean of the bracketed quantities over the subscript, and $\hat{\lambda}_{ij}$ is the ordinary least squares estimator of λ_{ij} calculated for a model that omits θ_i and δ_{ij}.

When the data do not include replicates (as is the case in the present data), this estimate is give by

$$\hat{\theta}_{ij'} = M_{j \# j'} [y_{ij} - \hat{\lambda}_{ij}].$$

Formal Definitions of Linear and Residual Accuracy

For simplicity, I present the definition of linear accuracy and residual accuracy for the case in which there are no replicates. In this case, the prediction of judge j regarding profile i could be expressed as $y_{ij} = \lambda_{ij} + \varepsilon_{ij}$, where λ_{ij} is as defined earlier and ε_{ij} is the residual from the linear model (note that in terms of the earlier notations, $\varepsilon_{ij} = \theta_i + \delta_i + \in_{ij}$). If φ_i is the value of the criterion associated with profile i, then the linear accuracy of judge j is the correlation between φ_i and λ_{ij}, and the residual accuracy of judge j is the correlation between φ_i and ε_{ij}.