Organizational Research Methods
Volume 12 Number 3
July 2009 403-417

10.1177/1094428107309326
http://orm.sagepub.com
hosted at
http://online.sagepub.com

Using Artificial Neural Networks to Model Nonlinearity

The Case of the Job Satisfaction–Job Performance Relationship

Mark John Somers
Jose C. Casal
New Jersey Institute of Technology

Neural networks are advanced pattern recognition algorithms capable of extracting complex, nonlinear relationships among variables. This study examines those capabilities by modeling nonlinearities in the job satisfaction–job performance relationship with multilayer perceptron and radial basis function neural networks. A framework for studying nonlinear relationships with neural networks is offered. It is implemented using the job satisfaction–job performance relationship with results indicative of pervasive patterns of nonlinearity.

Keywords: *job performance; job satisfaction; multiple regression; neural networks*

There has been increasing interest in the use of artificial neural networks (ANNs) in business research. Early applications of neural networks to business problems were mostly in finance and operations research in targeted areas such as credit scoring and bankruptcy prediction (cf. Vellido, Lisboa, & Vaughn, 1999). The thrust of this research was to establish the predictive efficacy of neural networks relative to conventional statistical methods and to explore the nuances of a new methodology with respect to data and problems encountered in business research.

Although comparatively fewer in number, studies using neural networks are also evident in organizational research. Most are in the area of organizational behavior with subareas including human resources management and vocational behavior. For example, ANNs were used to model employee turnover (Somers, 1999) and to study the relationship between work attitudes and job performance (Somers, 2001). Carson et al. (1999) studied vocational choices with ANNs whereas Palocsay and White (2004) used neural networks to study cross-cultural perceptions of justice in organizations.

These studies used the same approach as their counterparts in finance and operations research in that the primary purpose was to examine the predictive efficacy of ANNs relative to conventional statistical methods. As was the case in general business research, when used in organizational research, ANNs typically outperformed their more conventional counterparts. As such, it seems reasonable to conclude that neural networks perform at least as well and usually outperform linear statistical methods (cf. Palocsay & White, 2004; Scarborough & Somers, 2006).

Authors' Note: We wish to thank three anonymous reviewers for their guidance and contributions to this paper.

Work on neural computing has advanced to the point where it is clear when and why neural networks outperform multiple regression and related techniques. First, ANNs are nonparametric and, thus, are not bound by the assumptions of the general linear model (DeTienne, DeTienne, & Joshi, 2003). Second, neural networks model linear and nonlinear relationships among variables (Scarborough & Somers, 2006). Finally, ANNs are adaptive and "learn" patterns in data so that as new cases and datasets are collected, models adjust and remain generalizable (DeTienne et al., 2003).

Research has, therefore, established that ANNs are useful in estimating the upper boundary (e.g., ceiling) of explained variance when nonlinearity is present. Observed improvements in predictive efficacy, in turn, have been used to make a case for greater use of ANNs in organizational research (cf. Carson et al., 1999; Collins & Clark, 1993, Palocsay & White, 2004; Somers, 1999).

Although this research was a necessary first step in evaluating ANNs, improved predictive accuracy is not sufficient to justify wider use of neural networks in organizational research. Indeed, it is necessary to define a role for neural networks that includes a clear tieback to theory development and testing (Scarborough & Somers, 2006). This is a critical issue because neural networks are exploratory pattern-recognition algorithms so that it is critical to assess the plausibility and relevance of the relationships that they extract.

This general approach to using ANNs requires researchers to think differently about data analysis. Most important, it requires them to examine long–standing assumptions about the nature of the relationships among variables including the assumption of linearity. Thus, we begin the discussion of ANNs with the issue of nonlinearity in organizational research and then move on to consider how neural networks can be useful in modeling nonlinear relationships.

Nonlinearity, Neural Networks, and Organizational Research

Within the past decade, concerns have been expressed about the pervasiveness of linear thinking and linear methods in organizational research. Specifically, Starbuck and Mezias (1996) suggested that "researchers should consider using techniques other than squared-error regression, analysis of variance, and LISREL" (p. 115) when studying managerial perceptions to move beyond linear thinking. Bettis and Prahalad (1995) suggested that system dynamics related to strategy and adaptation are often nonlinear and need to be studied with nonlinear methods. Within the area of organizational behavior, Somers (1999, 2001) suggested that key relationships among variables might not be linear and called for greater use of nonlinear methods.

Although there is some agreement about the desirability of exploring nonlinear relationships, there is some confusion about the methodologies needed to accomplish this objective. Given the pervasiveness of linear thinking in organizational research, it is often difficult to specify the form of nonlinear relationships among variables in advance, as little guidance from theory or empirical findings is available. Thus, the nonlinear hypothesis is usually more general and is offered as an alternative to the implicit view that relationships among variables are linear.

Scarborough and Somers (2006) have considered this issue in detail and offerred a framework for assessing nonlinearity with neural networks. Studies can be either confirmatory or exploratory. The purpose of a confirmatory study is to formally test the assumption of linearity by using ANNs to establish that relationships among variables are primarily linear. Confirmatory studies are suited to domain areas where theory and research findings are well established and good reasons to think that a linear model is the best representation of reality are present.

In contrast, many areas in organizational research where exploratory studies of nonlinearity are appropriate are apparent. They are characterized by contradictory findings, weak empirical relationships in relation to theory-based predictions, and confusion about why progress is limited. In these cases, modifications to theory are usually unsuccessful in that new variables, and new processes produce little or no improvement in empirical results; that is, incremental variables do not result in incremental variance.

Often the one unexamined possibility, even after years of study, is that relationships among variables are not linear. For certain research topics, this omission becomes a concern, and the utility of exploratory studies assessing nonlinear relationships among variables becomes apparent. The impetus might be a specific concern that a linear model is inadequate for the phenomena under study (cf. Bettis & Prahalad, 1995), or it may be a more general notion that the form of the relationships of the variables under study is not well understood and the possibility of nonlinearity should be examined (cf. Guion, 1992).

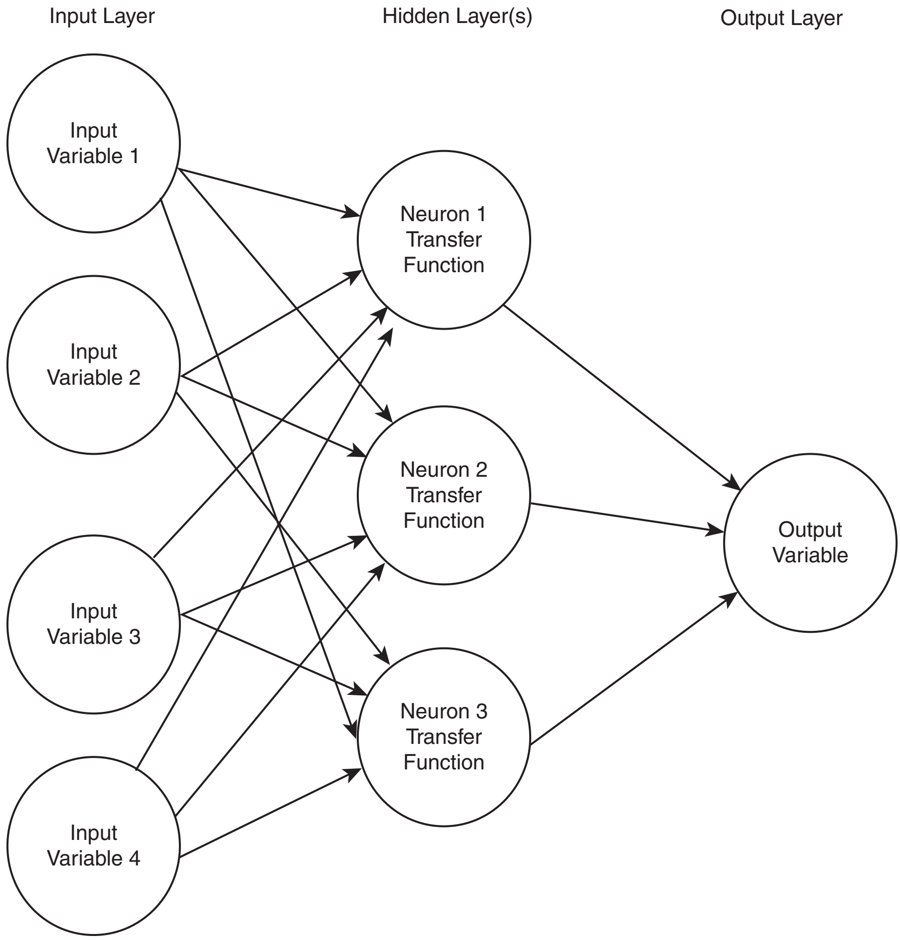
Modeling Nonlinearity With Neural Networks

Work on neural networks has advanced considerably, and many excellent books and journal articles that explain how ANNs identify patterns in data are evident in the literature (cf. DeTienne et al., 2003; Garson, 1988; Hornik, Strinchcombe, & White, 1989: Scarborough & Somers, 2006, White, 1990). There is no need to repeat this work here. Rather, our intention is to draw from it with specific emphasis on how neural networks capture nonlinear relationships among variables and how the form of those relationships is expressed.

Starting with a general definition, neural networks are statistical algorithms that capture patterns in data. More specifically, ANNs are pattern-recognition algorithms that capture salient features from a set of inputs and map them to outputs (cf. Swingler, 1996). To understand how mapping occurs, it is necessary to understand the statistical basis by which neural networks model patterns in data. Our discussion is restricted to supervised neural networks, that is, cases where there is a known target or criterion variable (cf. Bishop, 1995). These networks have the same basic architecture that includes an input layer, hidden layer(s), and an output layer (see Figure 1).

As indicated in Figure 1, each layer comprises a predetermined number of "neurons" or processing elements, which, in totality, define the architecture of a neural network. Using conventional terminology, the input layer represents the independent variables in the study whereas the output layer represents the dependent variable. Mapping occurs in the hidden layer where the number of neurons is discretionary.

Figure 1
A Neural Network Architecture With One Hidden Layer



Note: The input layer represents predictor or independent variables, and the output layer represents the dependent or criterion variable. Neural networks are built and trained to find patterns in data that can be mapped to the output variable. Mapping occurs in the hidden layer(s) of the network where input data are summed and weighted using statistical transfer functions. Predictors derived from the network are then compared with actual values from the data set, error is calculated, and results are transferred to the output layer usually with a linear transfer function. As data pass through the network many times, weights are adjusted and error is reduced. When statistical transfer functions are nonlinear (e.g., sigmoid), the neural network is capable of mapping complex patterns of nonlinearity.

The primary difference between neural networks and conventional statistical methods is that ANNs are adaptive (cf. DeTienne et al., 2003). That is, data are passed through the network many times such that each pass of data results in a predicted value that is compared to a known outcome (e.g., Y vs. Y hat). Adjustments are made to reduce error, and data are passed through the network until an acceptable reduction in error is attained.

This process is referred to as learning because as data pass through the network, error is reduced. Learning occurs in the hidden layer where input data are summed and weighted with statistical functions to generate a predicted value that is then passed on to the output layer usually with a linear transfer function. Pattern recognition is refined as weights are adjusted with each pass of data through the network architecture. To model nonlinear relationships, the neurons in the hidden layer must use nonlinear statistical functions. Specifically, it has been shown that if summed calculated inputs are transformed into an output using a nonlinear statistical function, the result is a model with true nonlinear parameters (DeTienne et al., 2003; Grznar, Prasad, & Tata, 2007; Swingler, 1996).

This process is affected by the type of statistical transfer function used and the number of neurons in the hidden layer. Therefore, different neural network architectures and/or different transfer functions can produce markedly different results. Furthermore, it should be noted that neural networks are not restricted to one hidden layer, and instances where two or more hidden layers are better suited to mapping nonlinearities in data are common (DeTienne et al., 2003; Ripley, 1996). More specifically, one hidden layer might not be sufficient to model complex nonlinear relationships (cf. Ripley, 1996) so that two or more hidden layers are necessary to accurately model relationships among variables (Hertz, Krogh, & Palmer, 1991). It is the responsibility of the researcher to demonstrate that the number of hidden layers in an ANN is appropriate for the problem at hand.

Neural Networks, Nonlinearity, and Data Visualization

One area where progress in neural computing has lagged is assessing the relative influence of predictor variables on the criterion variable. With no analog to the standardized coefficients produced by ordinary least squares (OLS) regression, assessing the relative influence of predictor variables on the criterion is problematic. Termed *sensitivity analysis*, initial attempts at assessing the relative contribution of input variables involved partitioning weights across the input and hidden and output layers and expressing them as a relative percentage of the network's output for each predictor variable (Garson, 1991). More recent approaches to sensitivity analysis involve systematically removing input variables from the network and then quantifying decrements in predictive accuracy (Bishop, 1995).

These advancements in sensitivity analysis are useful in addressing concerns about neural networks operating as "black boxes" that generate output that cannot be explained (cf. Astin, 2000); however, they are not well suited to understanding nonlinear relationships among variables. The reason is that nonlinearity, by definition, is characterized by disproportional relationships among variables such that small changes in one can lead to large changes in others (West, 1980). The ensuing inflection points (e.g., areas of high sensitivity) must be represented visually to be completely understood. Thus, when nonlinearity is central to a study, it is far more important to know the form of relationships among variables and the ranges where small changes have large effects than it is to know which variable acted as the strongest predictor.

Scarborough and Somers (2006) recommended using two- and three-dimensional graphs derived from an ANN after training is completed to understand how input variables affect the output variable and to identify areas of high sensitivity. This approach generates visual models of the disproportional relationships characteristic of nonlinearity, provides an indication of how pervasive it is, and identifies ranges in study variables where transition points are present. It also addresses the issue of interpretability by assessing whether or not the ANN has extracted relationships that make sense conceptually.

The Study

The current study extends prior research assessing the predictive efficacy of neural networks by modeling potential nonlinear relationships to gain new insights into relationships among variables. The content area concerns the relationship between job satisfaction and job performance. It was chosen for several reasons. First, the relationship between job satisfaction and job performance is a good illustration of a case where conventional thinking and conventional methods have fared poorly in relation to theory-based expectations; that is, observed relationships between job satisfaction and job performance have been modest at best and far below theory-based expectations (Judge, Bono, Thoresen, & Patton, 2001). Second, these disappointing results have led some researchers to question the form of the relationship between job satisfaction and job performance suggesting that it might not be linear and that a new method is needed to study it (cf. Guion, 1992). Finally, there is some supporting empirical evidence using neural networks indicating the presence of nonlinearity in relationships between work attitudes and job performance (Somers, 2001).

Method

Sample

The sample comprised nurses and psychiatric technicians drawn from a university medical center located in the southern United States. It was 88% female with a mean age of 35.6 years and a mean organizational tenure of 54.6 months. A survey of work attitudes, completed during normal working hours, was distributed and collected on-site. Respondents were asked to provide identification numbers to allow access to their personnel files from which tenure and job performance data were gathered. Confidentially was assured to respondents, and after behavioral data were collected from employee records, all identifying information was deleted. The response rate was 50% yielding 176 usable responses.

Measures

Tenure was measured in months and was taken from employee personnel records. Organizational commitment was measured using Meyer and Allen's (1984) scales for affective and continuance commitment ($\alpha = .80$ and .70, respectively). *Affective commitment* is defined as an emotional attachment to an organization whereas *continuance commitment* reflects investments in an organization that accrue over time. Job search behavior was measured with a scale developed and validated by Kopelman, Rovenpor, and Millsap (1992) ($\alpha = .83$). The scale assesses intensity of job search and asks respondents to identify those search behaviors that they engaged in the past 6 months (e.g., read employment ads). A 5-item facet-free measure developed by Quinn and Staines (1979) was used to measure job satisfaction ($\alpha = .78$). Job performance data were taken from the organization's formal performance appraisal process and reflected supervisor ratings of employee job performance. Data were gathered directly from employee personnel records. Ratings were based on multiple dimensions of job performance derived from job analysis, and those ratings were combined to form a weighted average that was measured along a 5-point scale with *5* being the highest rating possible.

Analysis

Data were partitioned into training and test subsets. Training data were used to estimate weights used by the ANN to subsequently generate predicted outcomes. In a process similar to cross-validation, test data represent a holdout sample. Weights derived from training are applied to test data and the predictions are compared to known outputs (Bishop, 1995). A significant decrement in performance indicates that the network was "overtrained" and uncovered patterns unique to the data set. Following convention (Bishop, 1995), training data comprised 80% of the sample (134 cases) whereas 20% was used for test data (39 cases).

Neural network model selection. Two ANNs were used in this study: multilayer perceptrons (MLPs) and radial basis functions (RBFs). MLPs are the most commonly used neural network (Swingler, 1996) and are well represented in organizational research (cf. Somers, 1999, 2001; Palocsay & White, 2004). MLPs use nonlinear transfer functions in the hidden layer to map nonlinearities in data. Sigmoid transfer functions were used in the current study. These S-shaped functions are able to find patterns of nonlinearity that linear statistics such as regression analysis cannot model (cf. DeTienne et al., 2003).

MLPs use backpropogation of error to adjust weights to improve predictive accuracy (Rummelhardt, Hinton, & Williams, 1986). The algorithm operates by minimizing differences between predicted and observed values by calculating the observed error for each pass of data and then adjusting weights in the model to make it slightly smaller. As data pass through the network, the objective is to minimize the sum of squared errors for training data (cf. Werbos 1988).

An advanced algorithm, conjugate gradient descent, that performs significantly better than does backpropagation (cf. Ripley, 1996; Shepherd, 1997) was used in the current study. Rather than adjusting network weights for each input vector, the conjugate gradient descent algorithm estimates the average gradient of the error surface across all cases and adjusts weights once at the end of each full pass of data.

In addition, data were analyzed with RBF neural networks. Neurons in the hidden layer of an RBF network operate as RBFs to reduce the distance between an input vector and a known output value using Gaussian transfer functions, which have been shown to be capable of modeling complex nonlinear relationships (Ripley, 1996).

As the neurons in the hidden layer are RBFs, they comprise a center that represents a weighted vector of input values, a distance measure that determines how far a given vector is from the center, and an output that represents the distance between the stored center and an input vector. As data pass through an RBF neural network, the Gaussian transfer function sends higher values to the output layer when the distance between the vector and the center is small (Ripley, 1996). Error is reduced using a clustering algorithm that minimizes the distance between input values and center values such that the neurons in the hidden layer with the smallest observed distances have the highest activation levels with respect to the output layer (Neural Ware, 1996). RBF networks were trained using a k-means clustering algorithm, which is the preferred method (Bishop, 1995).

ANN training and model refinement. Neural networks were built, trained, and evaluated using STATISTICA Neural Networks for Windows (2003). This software package is widely used and offers a full complement of features for building and evaluating neural networks.

Unlike conventional statistical methods, several discretionary elements are involved in building and training neural networks. The first is determining the basic network architecture including the number of hidden layers and the number of neurons in each hidden layer. Although several writers have noted that one hidden layer is usually sufficient to model complex nonlinear patterns (Bishop, 1995; Ripley, 1996), others have argued for using ANNs with more than one hidden layer (Hertz et al., 1991). We believe that this an empirical question that is best addressed with experimentation so that neural networks with more than one hidden layer were evaluated in the current study.

A second critical discretionary issue is determining when to discontinue training; that is, to determine when the network has "settled" and passing additional data through it results in capturing patterns in data that do not generalize. Hard-and-fast rules about when to stop training have not been established and the process varies by the problem being modeled and the type of ANN being trained (Scarborough & Somers, 2006).

In the current study, training was discontinued when no discernable reduction in error was evident. Although this is a common stopping point for training, there is no guarantee that a global or a local minimum has been reached. Overtraining was assessed by comparing performance on training and test data.

Comparison with regression analyses. Results from ANNs were examined in relation to those from OLS (cf. Palocsay & White, 2004; Somers, 2001) and Tobit regression models. This is accomplished by comparing correlations between predicted and observed values for test and training data. Tobit regression was used because research indicates that supervisors tend to avoid using the lower end of the performance scale (Kane, Bernadin, & Peyrefille, 1995) resulting in censored data. Tobit analysis is a modification of regression analysis designed for censored data (Takeshi, 1984), and analyses were conducted with STATA for Macintosh.

Results

MLP and RBF Neural Networks

Analyzing data with a neural network is an iterative process that involves experimentation with different network architectures and training parameters (cf. DeTienne et al., 2003; Scarborough & Somers, 2006). This process was initiated by using the artificial intelligence tools in STATISTICA Neural Networks to build the best performing MLP and RBF neural networks. The best performing RBF network had 4 neurons in the hidden layer

Table 1
Comparison of Predictive Accuracy of Neural Networks

Neural Network	Neurons in Hidden Layer(s)	Training Data		Test Data	
		R	R^2	R	R^2
RBF	2	.221	.049	.091	.008
RBF	4	.414	.171	.174	.030
RBF	5	.268	.072	.143	.020
RBF	7	.382	.146	.162	.026
RBF	9	.449	.201	.164	.027
RBF	15	.544	.296	.163	.027
RBF	25	.723	.523	.036	.001
MLP	2	.177	.031	.104	.010
MLP	5	.471	.222	.150	.023
MLP	10	.629	.396	.387	.150
MLP	15	.774	.599	.064	.004
MLP	20	.568	.323	.087	.008
MLP	10/5	.317	.100	.125	.016
MLP	10/3	.346	.120	.052	.003
MLP	10/4/3	.203	.041	.088	.008

Notes: MLP = multilayer perceptron; RBF = radial basis function.

while the best performing MLP had 10 neurons in the hidden layer. (R & R^2 for the MLP were .629 & .396 and .387 & .150 for training and test data respectively; see Table 1).

Network architectures were then varied by systematically reducing and increasing the number of neurons in the hidden layer. As indicated in Table 1, results were similar for the MLP and RBF networks. Fewer neurons in the hidden layer hampered pattern recognition and reduced predictive accuracy. Networks with an increased number of neurons in the hidden layer captured patterns in data that did not generalize as evidenced by strong performance on training data that was not sustained on test data.

As MLP networks performed better than did RBF networks, follow-on analyses with MLPs were conducted to ensure that one hidden layer was sufficient to model nonlinear relationships among study variables. MLPs with two hidden layers performed significantly worse than did those with one hidden layer. Specifically, the best results generated with a two hidden-layer MLP (10 neurons in the first hidden layer and 3 in the second) produced an R^2 of .34 with training data and an R^2 of .05 with test data. MLPs with three hidden layers underperformed those with two hidden layers with the best performing three-layer MLP producing an R^2 of .22 with training data and .04 with test data (see Table 1).

OLS and Tobit Regression Models

OLS and Tobit regression models had lower levels of predictive accuracy than did the best performing neutral networks. As indicated in Table 2, R^2 for OLS regression was .038 with training data and .005 with test data.

Although supervisory-rated job performance was measured along a 5-point scale, only 20% of the sample received ratings below 3.0 (kurtosis = 12.19). Tobit regression was used

Table 2
Comparison of Results With Ordinary Least Squares (OLS) and Tobit Regression Models

	Training Data ($n = 135$)	Test Data ($n = 39$)
Analysis	R^2	R^2
OLS regression	.038	.005
Tobit regression	.129	.059
MLP 5:10:1	.396	.150

Note: MLP = multilayer perceptron.

to correct for the censored job performance variable and produced a pseudo-R^2 of .129 on training data and an R^2 of .059 on test data (see Table 2). Although not statistically significant ($\chi^2 = 11.24$, $p > .05$), the Tobit regression model generated greater predictive accuracy than did OLS regression but lagged the best performing ANN by a considerable margin (R^2 of .381 vs. .129 with training data and .144 vs. .059 with test data).

Visual Analyses

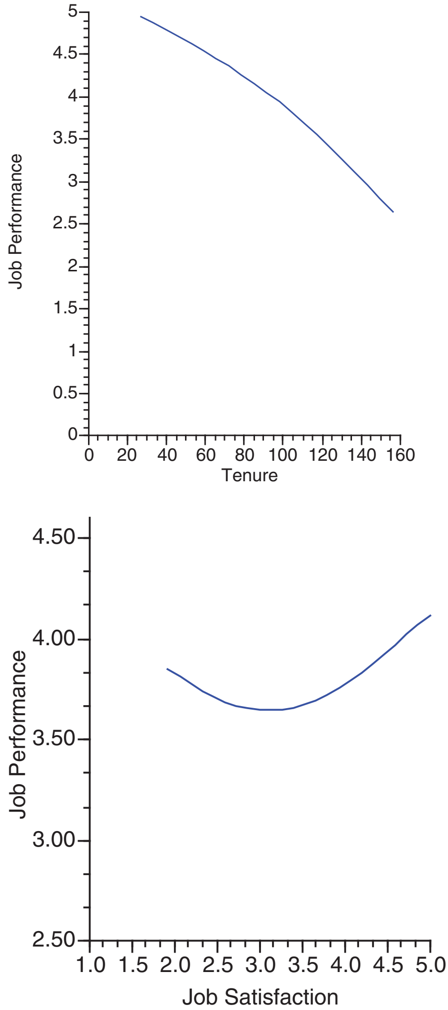
All visual analyses were conducted with the best performing neural network. Visual analyses represent patterns derived from training data mapped by a neural network at the completion of training. The Y axis, in turn, represents predicted values of job performance. In all cases, the graphs in Figures 2 and 3 represent relationships extracted by the MLP with all other variables in the model held constant at their mean levels.

Bivariate relationships between predictor variables and job performance are presented in Figure 1. As continuance commitment and tenure are tied to time-based investments in the organization, in the interest of brevity, only the relationship between tenure and job performance is presented. As indicated in Figure 2, all of the relationships between the predictor variables and job performance are linear except for job satisfaction that took the form of a *U* shape such that higher performance was associated with low and high levels of job satisfaction.

Triadic relationships are presented in Figure 3. As job satisfaction is the focal predictor in the current study and bivariate analyses indicated that is was the only variable that exhibited a nonlinear relationship with job performance, relationships between job satisfaction and job performance were explored in conjunction with continuance commitment, job search behavior, and tenure.

The analyses presented in Figure 3 are indicative of pervasive nonlinearity. The most complex pattern of nonlinearity includes job search behavior, job satisfaction, and job performance. The tilted *S* shape that defines the total response surface suggests that there is a relationship between satisfaction and performance at high levels of job satisfaction and low levels of job search; however, it is not as strong as that observed at the other extremes of these two predictor variables. Nonlinearity is also evident in the relationships between job satisfaction, continuance commitment and job performance, and job satisfaction, tenure, and job performance.

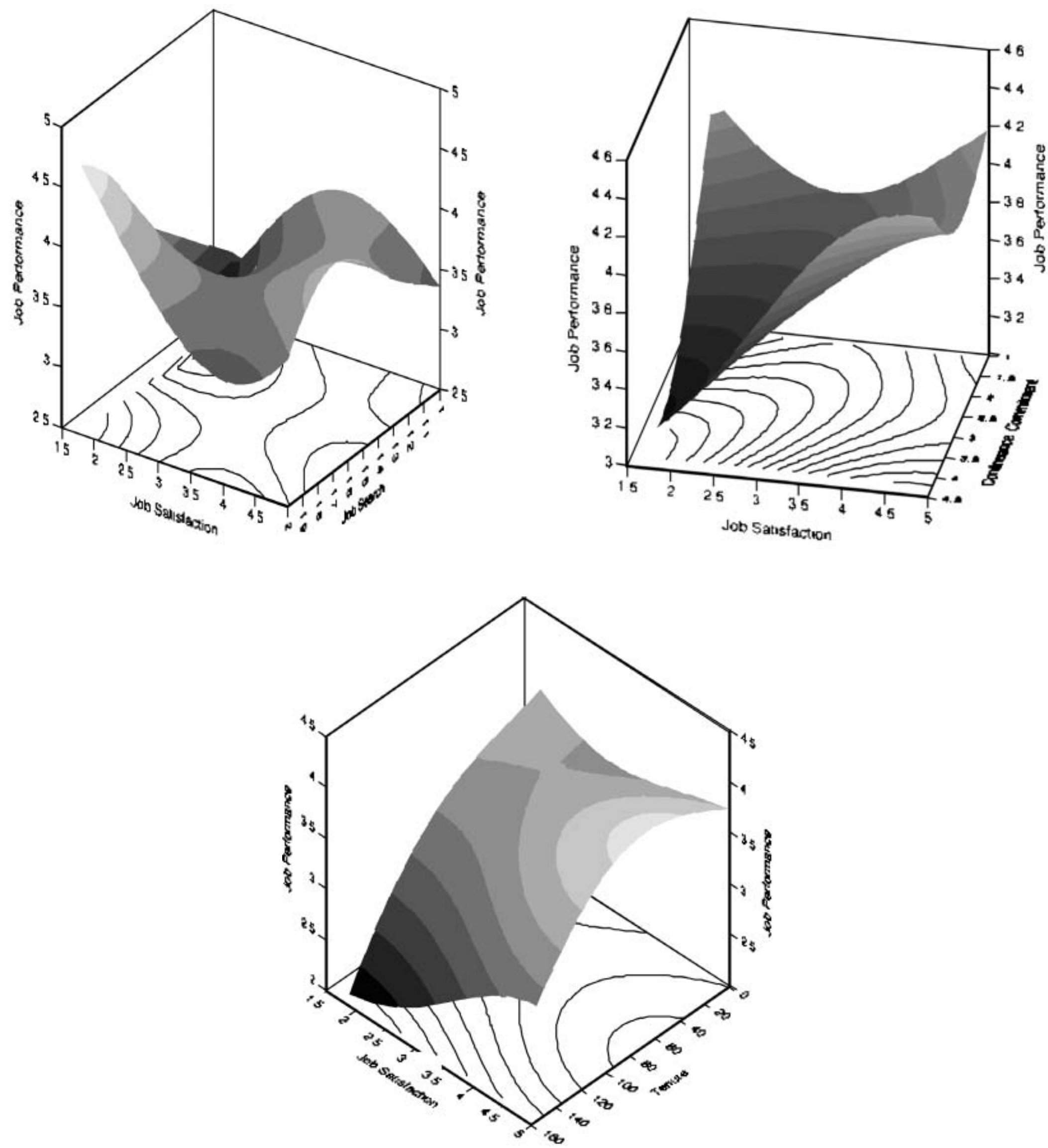
Figure 2
Form of Bivariate Relationships Between Predictor Variables and Job Performance Extracted by Multilayer Perceptron Neural Network



Discussion

DeTeinne et al. (2003) asked, "Why don't business researchers use neural networks?" (p. 236) and went on to point out that they are widely used in the hard sciences. The answer to this question a few years later is partly encouraging in that there is more interest in ANNs in organizational research and partly discouraging in that neural networks remain underutilized despite their potential to open up new areas of research and to provide new insights into areas where progress is limited (cf. Scarborough & Somers, 2006).

Figure 3
Nonlinear Response Surfaces of Job Satisfaction, Job Performance, Job Search, Continuance Commitment, and Tenure Extracted by Multilayer Perceptron Neural Network



One promising area for neural networks in organizational research is modeling nonlinearity. Although nonlinear models and methods are common in the physical and life sciences, most organizational research is grounded in linear thinking and linear methods. The extent to which this has hampered theory development and the robustness of empirical results is unclear; however, it has been suggested that domain areas such as strategy (Bettis & Prahalad, 1995), managerial perceptions (Starbuck & Mezias, 1996), leadership (Schneider & Somers, 2006), and work attitudes (Somers, 1999, 2001) might be advanced with nonlinear models and nonlinear methods.

ANNs are one such method, and though they offer the prospect of new insights into the form of relationships among variables, it is also imperative that their limitations are recognized. As neural networks are exploratory, it is important that conceptual justification for studying nonlinear relationships is present and that ANNs are one component of an integrated analysis plan that includes conventional statistical methods. It this regard, it is a misuse of ANNs to deploy them as the analysis of last resort in the hope of increasing explained variance, and then claiming that any observed increment, in and of itself, is meaningful (cf. Scarborough & Somers, 2006).

In the current study, several factors justified exploring a nonlinear model. First, there is a long history of poor empirical results leading calls for the form of the job satisfaction–job performance relationship to be studied using new methodologies (Guion, 1992). Second, the most recent meta-analysis (Judge et al., 2001) indicated that there was high interstudy variability that could not be accounted for by moderator effects suggesting that other factors might be present. Finally, one empirical study using ANNs (Somers, 2001) found evidence of nonlinearity that took the form of highly channeled relationships that, in turn, was offered as an explanation for the high interstudy variability observed by Judge et al. (2001).

The findings from the current study offer additional evidence of nonlinearity. Beginning with predictive accuracy, the best performing neural network explained more than twice as much variance in job performance than did Tobit regression with test data (.059 vs. .15) and significantly outperformed OLS regression (R^2 of .005 vs. .15). These results indicate that deviations from normality in the job performance variable are partly responsible for the poor results from OLS regression, but that nonlinearity is also present as ANNs produced much better results than did Tobit regression. They also demonstrate the importance of selecting benchmarking analyses that are useful in understanding results from neural networks.

Visual analyses offered insight into the form of nonlinear relationships among study variables. Bivariate analyses indicated that job satisfaction was the only variable that exhibited a nonlinear relationship with job performance. This finding might be a partial explanation for the modest corrected correlation between job satisfaction and job performance of .33 generated by meta-analysis (cf. Judge et al., 2001) and supports Guion's (1992) notion of a nonlinear satisfaction–performance relationship.

Mapping triadic relationships among variables produced complex, nonlinear response surfaces between job satisfaction, job performance, and contextual variables such as job search behavior even when relationships between contextual variables and job performance were linear. For example, visual analyses suggest that low levels of job satisfaction trigger job search among high performers that can be interpreted as overachievers becoming frustrated with their present situation and seeking change.

Findings such as these demonstrate the role of ANNs in theory development because they challenge researchers to explain areas of high sensitivity. That is, it is necessary to explain why the amplitude of a relationship changes so dramatically when a threshold value is crossed. In so doing, it opens up the possibility of new approaches to old problems. For example, current thinking ties decreasing satisfaction to low performance as a form of organizational withdrawal (cf. Judge et al., 2001), yet results from neural networks suggest that there is a underlying process whereby frustration reaches a threshold that triggers high performers to seek new jobs.

Theory development might lead to modification of existing processes or new conceptual frameworks that ultimately should translate into theory testing. To do so requires replicating patterns uncovered with ANNs with conventional statistical methods by building models with the appropriate higher order terms either using a cross validation sample or with new data (cf. Baatz, 1995). That is, it is necessary to move from exploratory to confirmatory studies that replicate patterns extracted by ANNs.

Given that ANNs efficiently model complex nonlinear relationships, they have the potential to open up new avenues of research and to provide new insights into long-standing problem areas. For this potential to be realized it is necessary for topic areas to be chosen carefully based on theory and prior research findings, for neural networks to be built and trained properly, and for analyses to go beyond increments in explained variance to offer insights into relationships among variables extracted by an ANN.

References

Astin, R. (2000). Neural networks: A science in trouble. *ACM SIGKDD Newsletter*, *1*, 33-38.

Barnard, E., & Wessels, L. (1992). Extrapolation and interpolation in neural network classifiers. *IEEE Control Systems, 12*, 50-53.

Baatz, E. (1995). Making brain waves. *CIO Executive, 9,* 23-29.

Bettis, R., & Prahalad, C. K. (1995). The dominant logic: Retrospective and extension. *Strategic Management Journal, 16*, 5-15.

Bishop, C. (1995). *Neural networks for pattern recognition*. New York: Oxford University Press.

Carson, A., Bizot, E., Hendershot, P., Barton, M., Garvin, M., & Kraemer, B. (1999). Modeling career counselor decisions with artificial neural networks: Predictions of fit across a comprehensive occupational map. *Journal of Vocational Behavior, 54*, 196-213.

Collins, J. M., & Clark, M. R. (1993). An application of the theory of neural computation to the prediction of workplace behavior: An illustration and assessment of network analysis. *Personnel Psychology, 46,* 503-524.

DeTienne, K. B., DeTienne, D. H., & Johsi, S. A. (2003). Neural networks as a statistical tool for business researchers. *Organizational Research Methods, 6*, 236-265.

Garson, D. (1988). *Neural networks: An introductory guide for social scientists*. London: Sage.

Garson, D. (1991). A comparison of neural network and expert systems algorithms with common multivariate procedures for analysis of social science data. *Social Science Computer Review, 9,* 399-434.

Guion, R. (1992). Agenda for research and action. In P. Smith (Ed.), *Job satisfaction: How people feel about their jobs and how it affects their performance* (pp. 257-281). New York: Lexington Books.

Grznar, J., Prasad, S., & Tata, J. (2007). Neural networks and organizational systems: Modeling nonlinear relationships. *European Journal of Operations Research, 181*, 939-955.

Hertz, J., Krogh, A., & Palmer, R. (1991). *Introduction to the theory of neural computation*. Reading, MA: Addison-Wesley.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward neural networks are universal approximators. *Neural Networks, 2,* 359-366.

Iaffaldano, M., & Muchinsky, P. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin, 97*, 251-273.

Judge, T., Bono, J., Thoresen, C., & Patton, G. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin, 127*, 376-407.

Kane, J., Bernadin, H. J., & Peyrefille, J. (1995). Stability of rater leniency: Three studies. *Academy of Management Journal, 38,* 1036-1051.

Kopelman, R., Rovenpor, J., & Millsap, R. (1992). Rationale and construct validity evidence for the Job Search Behavior Index: Because intentions (and New Year's resolutions) often come to naught. *Journal of Vocational Behavior, 40*, 269-287.

Meyer, J., & Allen, N. (1984). Testing the "side bet" theory of organizational commitment: Some methodological concerns. *Journal of Applied Psychology, 69,* 372-378.

Neural Ware, Inc. (1996). *Neural computing: A technology handbook for Professional II Plus and Neural Works Explorer*. Pittsburgh, PA: Author.

Palocsay, S. W., & White, M. M. (2004). Neural network modeling in cross-cultural research: A comparison with multiple regression. *Organizational Research Methods, 7,* 389-399.

Quinn, R., & Staines, G. (1979). *The 1977 Quality of Employment Survey.* Ann Arbor: University of Michigan, Institute for Social Research.

Ripley, B. (1996). *Pattern recognition and neural networks*. Cambridge, UK: Cambridge University Press.

Rummelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. E. Rummelhart, J. L. McClelland, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318-459) Cambridge, MA: MIT Press.

Scarborough, D., & Somers, M. J. (2006). *Neural networks in organizational research: Applying pattern recognition to the analysis of organizational behavior*. Washington, DC: APA Books.

Stat Soft, Inc. (2003). *Neural networks in STATISTICA 7.0*. Tulsa, OK: Author.

Schneider, M., & Somers, M. (2006). Organizations as complex adaptive systems: Implications of complexity theory for leadership research. *Leadership Quarterly, 17,* 351-365.

Shepherd, A. J. (1997). *Second-order methods for neural networks: Fast and reliable training methods for multi-layer perceptrons*. New York: Springer.

Somers, M. J. (1999). Application of two neural network paradigms to the study of voluntary employee turnover. *Journal of Applied Psychology, 84*, 177-185.

Somers, M. J. (2001). Thinking differently: Assessing nonlinearities in the relationship between work attitudes and job performance using a Bayesian neural network. *Journal of Occupational and Organizational Psychology, 74,* 47-61.

Starbuck, W. H., & Mezias, J. M. (1996). Opening Pandora's box: Studying the accuracy of managers' perceptions. *Journal of Organizational Behavior, 17*, 99-117.

Swingler, K. (1996). *Applying neural networks: a practical guide*. New York: Academic Press.

Takeshi, A. (1984). Tobit models: A survey. *Journal of Econometrics, 24*, 3-61.

Vellido, A., Lisboa, P. J. G., & Vaughn, J. (1999). Neural networks in business: A survey of applications (1992-1998). *Expert Systems with Applications, 17*, 51-70.

Werbos, P. (1988). Generalization of backpropogation with application to a recurrent gas market model. *Neural Networks, 3,* 535-549.

West, B. J. (1980). *Lecture notes in biomathematics: An essay on the importance of being nonlinear*. Berlin, Germany: Spring-Verlag.

White, H. (1990). Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings. *Neural Networks, 3,* 535-549.

Mark Somers is a professor of management at New Jersey Institute of Technology and a member of the graduate faculty of Rutgers University in Newark. He is interested in commitment processes in organizations, employee socialization and retention, and nonlinear modeling.

Jose Casal is a special lecturer at New Jersey Institute of Technology. His research interests include ethical conduct in organizations, public sector management, and neural computing.