

# Human decision-making behavior and modeling effects

Choong Nyoun Kim <sup>a,\*</sup>, Kyung Hoon Yang <sup>b</sup>, Jaekyung Kim <sup>c</sup>

<sup>a</sup> *Department of Management, University of Seoul, Dongdaemun-Gu Jeonnon-Dong 90 Seoul, South Korea*

<sup>b</sup> *Department of Information Systems, University of Wisconsin-La Crosse, 1725 State Street, La Crosse, WI, USA*

<sup>c</sup> *Department of Management, University of Nebraska–Lincoln, Lincoln, NE, USA*

Available online 23 June 2007

## Abstract

Previous research indicates that the human decision-making process is quite non-linear and that non-linear models would be more suitable than linear models for developing advanced decision-making models. In our study, we tested this generally held hypothesis by applying linear and non-linear models to experts' decision-making behavior and measuring the predictive accuracy (predictive validity) and valid non-linearity. As a result, we found that non-linearity in the decision-making process is positively related to the predictive validity of the decision. Secondly, in modeling the human decision-making process, we found that valid non-linearity is positively related to the predictive validity of non-linear models. Thirdly, we found that the more non-linearity is inherent in the decision-making process, the more non-linear models are effective. Therefore, we suggest that a preliminary analysis of the characteristics of expert decision-making is needed when knowledge-based models such as expert systems are being developed. We also verify that the lens model is effective in evaluating the predictive validity of human judgment and in analyzing the validity and non-linearity of the human decision-making process.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** Valid non-linearity and predictive validity in human decision-making behavior; Modeling effect; Predictive accuracy; Human judgment; Lens model analysis

## 1. Introduction

A great number of studies on human decision-making and judgment have been made in the field of social science, and a variety of methodologies have been researched [9]. Understanding the human decision-making process and the modeling of the decision-making process is one of the goals of this discipline [17]. Studies on decision-making can be classified into two

categories: the study of decision modeling and the study of decision process tracing [19,30].

Decision modeling studies the human decision-making mechanism and tries to build models that predict human decisions. This field has been researched under the name of expert systems in the discipline of management, and the findings are abundant [5,8,14,22,31]. Examples of findings include the development of new algorithms for building decision-making models and the development of methodologies for a knowledge base. In the past, studies were focused on modeling that resembled expert decisions and judgment. However, using enormous amounts of real data, recent studies have rigorously investigated the modeling of rules and associations. The application of these results has been

\* Corresponding author.

E-mail addresses: [cnkim27@uos.ac.kr](mailto:cnkim27@uos.ac.kr) (C.N. Kim),  
[yang.kyun@uwlax.edu](mailto:yang.kyun@uwlax.edu) (K.H. Yang), [jkim6@unlnotes.unl.edu](mailto:jkim6@unlnotes.unl.edu) (J. Kim).

expanded to a variety of areas, such as finance (e.g., bankruptcy prediction and stock price index prediction), marketing, account auditing, credit rating, and venture investment decision-making [33,34,36].

Decision process tracing, the other paradigm of study for human decision-making, focuses on the process of judgment and decision-making. To measure the predictive validity of human judgment, this research has introduced various methods for the analysis of the decision-making process [13] these methods include probability scoring rules, log transformation, and the lens model. In probability scoring rules, mean probability scores (MPS) are viewed as efficient tools for measuring the level of uncertainty [19,35]. Einhorn [11,12] suggested log transformation to classify types of human decision strategies. The lens model, proposed by Brunswick [6] and developed by Tucker [32], describes decision-making behavior in terms of linearity and non-linearity. This model also provides the tools to measure the predictive validity of the linear and non-linear parts of decision-making behavior. The effectiveness of the lens model has been verified by various empirical studies [17,19,28,37].

Decision modeling and decision process tracing are not independent because the research in both fields has a common goal which is to improve decision quality by analyzing and understanding human decision-making behaviors. Therefore, a combination of these two areas of research presents several meaningful issues. First, decision process tracing methods can be used in evaluating the predictive validity of decision-making models. Second, an extensive analysis of the decision process tracing methods can help develop an advanced decision-making model. Third, this combined approach may explain the reasons why non-linear models and statistical linear models show contradictory results for the same problem. Even though many researchers [19] insist that non-linear models, such as the neural network model, show better performance in some studies, much research still shows that statistical linear models are better in many fields. When we take into account the fact that the performance of a model depends on the input data as well as the model itself, it would be reasonable to include characteristics of the input data, in addition to the features of decision-making behaviors when evaluating the model's predictive validity. This inclusion would be helpful in explaining the contradictory results between previous behavioral accounting studies, which assert the superiority of the statistical linear model, and recent studies, which assert the superiority of the non-linear model. Also, this contradiction might have occurred, not because of the

model, but because of the input data. If the predictive validity of the decision-making model is affected by the level of the linearity/non-linearity of the problem and by the validity of the input data, this should be taken into account in the selection of model development techniques.

In respect to these issues, we analyzed decision-making, studied model building and evaluation, and investigated the relationship between the characteristics of decision-making behavior (non-linearity and its validity) and the predictive validity of the models.

## 2. Literature review

### 2.1. Decision process tracing

The analytical framework needed to understand the human decision-making process was borrowed from studies on human judgment in the cognitive psychology discipline [36]. The decision-making process is a major branch of decision-making studies. Finding the key factors affecting the decision-making process has been a core research topic of previous studies. As a consequence of this research, types and characteristics of decision-making behavior and measuring methods and/or models have been developed [13,24]. For instance, mean probability scores (MPS) were considered a useful method to measure predictive accuracy [19]. MPS is a function of squares of the deviation score between predicted values and actual outcome values. The formula of MPS is shown in Fig. 1.

MPS, which is an error measurement method, is widely used for assessing human predictive validity along with the hit ratio [19]. The value of MPS is between 0 and 1, and “MPS=0” means that all predictions are accurate, while “MPS=1” means that all predictions are inaccurate. For example, if a certified public accountant (CPA) predicted the possibility of bankruptcy for two banks, A and B, with the possibility 0.7 and 0.6 respectively, and they actually bankrupted later, the hit ratio is 100%, and the MPS is 0.125  $\left(\frac{[(1-0.7)^2 + (1-0.6)^2]}{2}\right)$ .

Einhorn [11], [12] assumed that human decision-making behavior is non-linear rather than linear.

$$MPS = \frac{\sum(\text{outcome-prediction})^2}{N}$$

where

Outcome: actual result with values 0 or 1

Prediction: probabilistic prediction with values between 0 and 1

Fig. 1. Mean probability score.

Furthermore, they provided methods for classifying human decision-making behavior into two types conjunctive and disjunctive. This classification has been verified by several researchers [18].

First proposed by Brunswick [6], the lens model has been developed by others [15,32] to investigate the use of non-linearity in human decision-making behavior. Many behavioral accounting researchers have discussed the use of the lens model with regard to the examination of a human judgment [21]. A detailed description of the lens model appears in Kim and McLeod [17].

## 2.2. Research on the decision-making model

Since the 1960s, research on decision-making has been explored in accounting and management sciences, as well as in many other fields. Main concerns were how to mimic human decision-making and determine if a model predicts better than human experts do. This is why the research in this field was called “Judgmental Bootstrapping.” A number of empirical results support the value of judgmental bootstrapping [23]. Key modeling methods are linear, such as linear regression analysis and discriminant analysis. Although there is research that shows otherwise [8,20,27], most experimental studies found that linear models predict the actual outcome more accurately than human experts do [10,19]. Previous researchers in this field state that these kinds of positive results are due to the simple linearity of the model, which reduces the judgmental inconsistency of human decision-making behavior [17,29].

Since the late 1970s, studies on decision-making models in the field of management have continued with expert system and artificial intelligence research. Several modeling methodologies and algorithms have been researched to extract expert knowledge and decision-making behavior. One of these methodologies is the inductive learning approach, which treats input data as non-linear. Quinlan’s [26] ID3 (also evolved to C5 later) is the most widely used algorithm in inductive learning approaches. ID3 represents the human decision process as a tree-structured model and shows a very prominent prediction accuracy compared to traditional statistical approaches. ID3 shows a higher prediction accuracy, especially when sample data are stable with less noise [17]. In its earlier stage of development, ID3 could only treat discrete data and make a binary classification; however, as it has evolved, it can also handle continuous data, making more sophisticated classifications possible. The other approach is neural network, which is broadly applied in modeling management decision-making. Neural network models are widely

used in a variety of applications because they are free from statistical assumptions, making it easy to find non-linear relationships among input and output variables. Furthermore, they show better performance in dealing with noisy sample data.

## 2.3. Comparative studies for the linear and non-linear models

The decision-making model is mainly applied to classification and/or prediction problems. Most classification researchers have used hit ratio for the performance evaluation criterion. They have also used statistical models, such as regression analysis, discriminant analysis, and logistic analysis, which are based on linear relationships among variables [8]. These statistical models have been used mainly to analyze and model expert decision-making behavior in behavioral sciences such as psychology and behavioral accounting, where they have proved their prowess [3,7,10]. Non-linear models, such as neural network and tree structure have been evaluated by comparing their results with those of statistical linear models, which have been considered reliable [8].

However, Chung and Silver [8] argued that the comparison of non-linear and linear models was only based on input data without considering the type of tasks or decision-making behavior to which the methods were applied. Previous research on expert systems also committed these kinds of mistakes without considering circumstantial factors such as the characteristics of input data and the expert’s behavior in the research model design. Characteristics of data or those of an algorithm in models may also distort the model performance. For example, the degree of linearity and non-linearity of input data can distort the performance of a model when a linear model analyzes the non-linearity of input data or a non-linear model analyzes the linearity of input data.

Therefore, a comparison study should consider both the linearity/non-linearity of the input data and the linearity/non-linearity of the model. An analysis of environmental factors, such as the characteristics of input data and the participating expert’s behavior, should come first; and the result of the preliminary analysis should be used in the performance evaluation of models.

There is a second issue. In previous research, hit ratio was popularly used as a performance index in expert systems. In this case, much information can be lost since the decision is always “0” or “1”. Therefore, the hit ratio may not be appropriate as a performance measurement because it does not reflect the level of uncertainty and/or competence of the expert’s knowledge, experience, or

judgment. To overcome this limitation, an additional performance index is required to measure the predictive validity more accurately.

Consequently, it is believed that when evaluating model performance, the model should include environmental factors such as the characteristics of input data and the participating expert's behavior, as well as the uncertainty and competence of the expert's knowledge, experience, or judgment.

#### 2.4. Research objectives

It seems that a combination of research on decision-making modeling and on decision process tracing would be valuable not only theoretically but also practically, as Svenson [30] has insisted. But, despite his intention to that effect, there has been a lack of effort put into this kind of research. In this study, we analyze the findings and methods of both types of research and combine them to overcome the weakness of each. Through our analysis, we expect to find a relationship between characteristics of decision-making behavior and modeling methods. This study focuses on human decision-making behavior and the validity of the decision-making behavior. More specifically, linear and non-linear behavior is differentiated, and the validity of non-linear behavior is examined. To do this, the *C* index of lens model analysis is used, which is intended to represent the valid non-linearity in decision-making behavior [3,17,19].

This study has three objectives. First, we analyze the experts' decision-making strategies in terms of linearity, non-linearity, and validity of non-linearity with the *C* index of lens model analysis. A high *C* index value indicates the existence of valid non-linearity in decision-making behavior, which contributes to the predictive validity of the human decision maker. Second, predictive models for each human subject are built based upon one linear and two non-linear algorithms, and model performance is compared in terms of MPS as well as hit ratio. The correlations between the predictive validity of models and the *C* index of each subject are examined. Third, this study analyzes the argument of the previous research [17,20] concerning the relationship between the model performance and the human experts: "The more valid (accurate) human subject modeled, the less bootstrapping by linear models is likely because linear models cannot capture the valid non-linear decision behavior of the human subject." We seek to find which type of model is more valid for prediction when valid non-linearity is inherent in decision-making behavior. This approach may reveal that the conflicting results of previous comparative studies [7,8,10,18,19,22] between

statistical linear models and non-linear models were caused by the non-linear characteristics of decision-making behavior or the non-linear properties of the input data.

#### 2.5. Research model and hypotheses

Two hypotheses concerned with the objectives were developed.

**Hypothesis 1.** *Valid non-linearity of decision-making is positively related to the predictive accuracy of a non-linear model, but not positively related to the predictive accuracy of a linear model.*

This hypothesis is based on the presumption that human decision-making is basically non-linear [2,4,11,18,20,21], and the validity of decision-making is decided by valid non-linearity [12,20]. Validity means the predictive accuracy of decision-making, and valid non-linearity is defined as the non-linear portion that affects the accuracy of decision-making. We expect that, since valid non-linearity of decision-making is explained better by a non-linear model, a non-linear model leads to better predictive validity than a linear model if there is valid non-linearity decision-making behavior. To examine the valid non-linearity in the subjects' decision-making behavior, *C* index is a useful and general measure [17–19,32]. Since the valid non-linearity should be captured more successfully by non-linear algorithms, a high *C* index value should be correlated with the predictive validity of non-linear models.

**Hypothesis 2.** *As the validity of a subject's decision-making increases, the modeling effect of a linear model significantly decreases to a greater extent than the modeling effect of a non-linear model.*

Hypothesis 2 is concerned with Libby's argument [20], which asserts the limitation of linear algorithms in simulating or modeling human experts' decision-making behavior. To test the argument [20], the concept of modeling effect is used [17,20]. The modeling effect is defined as the improvement of the predictive validity by modeling the human decision-making behavior. It is expressed as the incremental accuracy of the model over the accuracy of a human subject (computed by subtracting the validity of a human subject from the accuracy of the model of the human subject) [17].

By testing the two hypotheses, we can explore the relationship between the model's predictive validity and

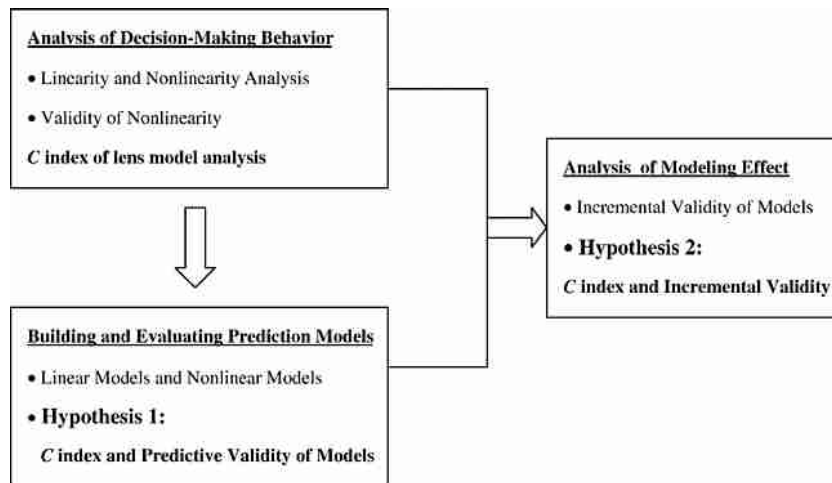


Fig. 2. Research model.

decision-maker's behavioral characteristics. Fig. 2 depicts our research model and hypotheses.

### 3. Research method

#### 3.1. Task: bankruptcy prediction

We chose bankruptcy prediction as the experimental task to test our research hypotheses. Bankruptcy prediction has been one of the most frequently studied human decision-making tasks since Altman's research [1]. As a result, we can compare our study to many previous studies. Bankruptcy prediction is also directly related to many applications, such as credit ratings, bank loans, and venture investment decisions [17,22,31,36].

#### 3.2. Data and participants

Our data was obtained from thirty bankrupted and thirty non-bankrupted companies in the U.S. in 1985. To maintain the consistency of data quality, we extracted sample companies from the same industry (manufacturing) with a similar size of about \$50 million in average assets. We used financial data for the two years prior to bankruptcy for each company. We kept the names of the companies anonymous. The ten most frequently used financial ratios from previous studies [16,17] were considered: 1) net income/total assets (profitability), 2) current assets/sales (Activity), 3) current assets/current liability (liquidity ratio), 4) current assets/total assets (asset balance), 5) cash/total assets (cash position), 6) total debt/total assets (financial leverage), 7) [current assets – current liability]/total assets (relative working capital), 8) sales/total assets (sales-generating ability of assets), 9) retained

earnings/total assets (cumulative profitability), 10) [current assets/current liability]/sales (working capital turnover).

Participants were selected from two groups: One group consisted of 16 experts who worked as certified public accountants (CPAs) or as financial CEOs who have CPA experience; the other consisted of 24 graduate students majoring in finance and accounting. To increase reliability, we adopted the test–retest approach. First, participants were asked to predict the bankruptcy/no-bankruptcy status of 70 cases. They were asked 10 cases twice, for a total 60 different cases. The participants whose prediction rate for the ten duplicated cases was lower than 80% consistent were eliminated. Eight students were eliminated, and 32 participants were selected. We expected that the prediction accuracy and the decision-making behaviors of the two groups would be different.

#### 3.3. Experiment procedure

The experiment in this study consists of three steps. In the first step, each participant predicts the bankruptcy possibility of 60 sample companies. Each participant makes two types of predictions: a binary decision and a probability prediction. In the binary decision, each company is labeled as either “0” (bankruptcy) or “1” (no-bankruptcy). In the probability prediction, each company is labeled by a ten-level, quasi-continuous scale based on the participant's confidence in the decision. If a participant predicts bankruptcy for a company, he/she may choose from “0.0” to “0.4,” where “0.0” implies the highest confidence and “0.4” the lowest confidence of bankruptcy.

If a participant predicts no-bankruptcy, he/she chooses a value from “0.6” to “1.0,” where “1.0” means the highest confidence and “0.6” the lowest confidence of no-bankruptcy. The reason that we use both the binary decision and the probability prediction is that certain types of algorithms cannot be fairly evaluated if the object variable is considered as either discrete or continuous.

In the second step, the prediction performance of each participant is evaluated by the hit ratio and MPS. The portion of non-linearity in the decision-making and the validity of non-linearity are also examined by the lens model’s *C* index and  $R_a$  [17,23]. The lens model consists of two types of linear models. One is the linear regression model of a participant’s bankruptcy prediction and the independent variables of ten financial ratios, and the other is the regression model of the actual result of bankruptcy and the independent variables. The quasi-continuous value of a participant’s prediction confidence was used as a dependent variable in the regression model.

In the third step, prediction models are developed based on each participant’s prediction. Statistical linear regression is adopted for the linear model, and Quinlan’s [26] C4.5 and the back-propagation paradigm of neural network are adopted for the non-linear model. It is known that the number of hidden layers is positively related to the overfitting of training, and it is recommended to use fewer hidden layers than the number of input nodes [25] in building the neural network model. Though there is no rule for the exact number of hidden layers, many previous studies used one hidden layer [17]. In our research, the neural network model consists of ten input nodes, five hidden nodes in one hidden layer, and one output node. The sigmoid function was used for the transfer function, and the delta rule was used for the learning algorithm. We repeatedly used this network model 32 times for each participant. The participant’s prediction confidence level was used as the objective variable in the training samples, while actual bankruptcy was used in the testing samples. This makes it possible to divide training samples and testing samples and to use both hit ratio and MPS for the measurement of predictive validity of linear or non-linear models.

## 4. Analysis and results

### 4.1. Analysis of decision-making

The prediction accuracies of 32 participants were measured by hit ratio, MPS, and  $R_a$  of the lens model.

Table 1

Prediction accuracy analysis of participants

Participant	$R_a$	Hit ratio (%)	MPS	<i>C</i> index	<i>C</i> index ( <i>t</i> -statistics)
1	0.552	78	0.1767	0.245	**1.92451
2	0.630	80	0.1548	0.528	**4.73494
3	0.657	83	0.1430	0.534	**4.81005
4	0.573	77	0.1757	0.287	**2.28171
5	0.579	75	0.1708	0.236	**1.84956
6	0.511	73	0.2057	0.244	**1.91616
7	0.546	78	0.1758	0.186	*1.44169
8	0.447	73	0.2105	0.104	0.79635
9	0.584	80	0.1655	0.310	**2.48322
10	0.426	68	0.2180	0.026	0.19807
11	0.489	70	0.2115	0.289	**2.299061
12	0.517	75	0.1963	0.15	1.155439
13	0.609	77	0.1708	0.359	**2.929339
14	0.496	75	0.1962	0.203	*1.578876
15	0.399	70	0.232	−0.051	−0.38891
16	0.517	68	0.2048	0.281	**2.229879
17	0.558	72	0.1802	0.209	*1.627642
18	0.496	73	0.1933	0.048	0.365979
19	0.488	72	0.1932	0.157	1.210691
20	0.363	65	0.2648	0.154	1.186989
21	0.532	70	0.1948	0.259	**2.04217
22	0.569	73	0.1732	0.201	*1.562662
23	0.483	70	0.2083	0.253	**1.991584
24	0.394	70	0.2553	0.057	0.434806
25	0.307	63	0.261	0.004	0.030463
26	0.342	63	0.2277	−0.101	−0.77315
27	0.425	68	0.2473	0.051	0.388911
28	0.444	70	0.2072	−0.097	−0.74223
29	0.338	67	0.2643	0.202	*1.570767
30	0.363	68	0.2587	0.157	1.210691
31	0.249	62	0.2768	−0.111	−0.85061
32	0.348	72	0.2475	−0.001	−0.00762

\*  $\alpha < 0.1$ , \*\*  $\alpha < 0.05$ .

The valid non-linearity of a participant’s decision-making behavior was measured with *C* index of the lens model. Hit ratios for most of the participants are around 70%, and the highest is 83% (participant # 3). Most MPS measures are around 0.2. Values of  $R_a$  are greater than the values of previous studies [17]. We assume the reason for this is because the predictive variable is not discretely measured, but rather quasi-continuously measured, with ten-level prediction confidence. *C* index shows the validity of the non-linearity of 12 participants’ decisions, with a significant level of  $\alpha = .05$ , and of 5 participants’ decisions, with a significant level of  $\alpha = 0.1$ . The results of the prediction accuracy of 32 participants are summarized in Table 1.

Table 2 shows the correlation among  $R_a$ , hit ratio, MPS, and *C* index. The correlation coefficient between  $R_a$  and hit ratio is relatively high at 0.8701. This means that  $R_a$ , which is used to measure predictive validity in the lens model, might be an appropriate measurement

Table 2  
Correlation between *C* index and prediction accuracy

	Hit ratio	MPS	<i>C</i> index ( <i>p</i> -value)
$R_a$	0.8701	−0.8598	0.8026 ( $p < 0.0001$ )
Hit ratio	–	−0.8598	0.7129 ( $p < 0.0001$ )
MPS	–	–	−0.6889 ( $p < 0.0001$ )

for prediction accuracy. Correlation coefficients between  $R_a$  and MPS and between hit ratio and MPS both are −0.8598. This high correlation may come from the fact that MPS also measures prediction accuracy. MPS shows the negative relationships with other indexes because it uses the prediction error that caused the negative sign.

Correlation coefficients between *C* index (measuring valid non-linearity) and prediction accuracy measurements are high: the  $R_a$  is 0.8026, the hit ratio is 0.7129, and the MPS is −0.6889. They are also statistically significant at  $\alpha = .01$  level. Therefore, as expected, we can conclude that valid non-linearity is highly correlated with the validity of decision-making.

These results support the findings of previous studies [19]: Valid non-linearity is an important factor contributing to predictive validity. These results also mean that valid non-linearity can be used to measure predictive validity. To examine the value of *C* index as an evaluation index of prediction accuracy, we classified the samples as valid (with 0.2 or higher *C* index value) and invalid (less than 0.2) and compared the two groups' prediction accuracy. The two groups have statistically different ( $p$ -value  $< 0.001$ ) prediction accuracies in terms of  $R_a$ , hit ratio, and MPS; and this result shows that valid non-linearity-based classification is consistent with the classification based on predictive accuracy. Table 3 shows the results in detail.

#### 4.2. Model construction and evaluation

The decision-making models applied to 32 participants were developed using a linear regression model, a

Table 3  
Prediction accuracy of participants (group average)

Group	<i>C</i> index	$R_a$	Hit ratio	MPS
Valid group ( $n = 16$ , <i>C</i> index $> 0.2$ )	0.290	0.542	74.25%	0.187
Invalid group ( $n = 16$ , <i>C</i> index $< 0.2$ )	0.046	0.409	69.38%	0.229
<i>t</i> -statistics ( <i>p</i> -value)	6.781 ( $p < 0.001$ )	4.865 ( $p < 0.001$ )	2.998 ( $p < 0.001$ )	3.980 ( $p < 0.001$ )

Table 4  
Group performance comparison among prediction models

Group	Linear regression model		Tree structure model		Neural network model	
	Hit ratio	MPS	Hit ratio	MPS	Hit ratio	MPS
Valid group ( $n = 16$ , <i>C</i> index $> 0.2$ )	61.14%	0.195	71.46%	0.187	73.23%	0.187
Invalid group ( $n = 16$ , <i>C</i> index $< 0.2$ )	57.08%	0.201	67.10%	0.210	65.94%	0.228
<i>t</i> -value	2.206*	−0.88	1.697*	−2.34*	4.883*	−3.95*

\* $\alpha < 0.05$ .

tree structure model using C4.5, and a neural network model using back propagation. Hit ratio and MPS are used as evaluation criteria. In the linear regression and neural network models, we used the value 0.4 as a threshold to evaluate hit ratio because prediction values are continuous values in these two models. For consistency purposes, this threshold is also used for the tree structure model. It is reasonable that the value 0.4 is counted as a threshold because we classified the non-bankrupt company with a prediction confidence of greater than 0.6 and the bankrupt company with less than 0.4 in the training sample. 0–0.4 means bankruptcy and 0.6–1.0 means non-bankruptcy. In the previous researches, the tree structure model may have been evaluated more unfairly than the linear regression model or neural network model when measuring MPS because its objective variables is discrete, while the objective variable of the linear regression model and neural network model is continuous. However, in this study, we calculated MPS under the assumption that the objective variable of the tree structure model is continuous because the objective variable is measured by quasi-continuous scale values (0.0–1.0).

In Table 4, we classified participants into groups of valid and invalid based on the *C* index and analyzed the two groups by linear regression, tree structure, and neural network. After that, we calculated the average prediction accuracy of the models for each group. Since the purpose of this study is not model comparison but the investigation of the effect of valid non-linearity in modeling, we skipped the model performance comparison. The difference between groups indicates that valid non-linearity of decision-making behavior is directly related to the predictive validity of the model. The results of the two groups, which are classified by *C* index, show significant differences in the hit ratio and MPS between the tree structure model and the neural

Table 5

Top ten ranking of prediction accuracy using hit ratio

Participant/Model	Hit ratio (%)	Rank
Tree structure model (Participant 3)	83.3	1
Participant 3 ( <i>C</i> index: 0.534)	83.0	2
(1st ranked participant in prediction accuracy)		
Neural network model (Participant 9)	80.0	3
Tree structure model (Participant 9)	80.0	3
Participant 9 ( <i>C</i> index: 0.310)	80.0	3
(2nd ranked participant in prediction accuracy)		
Participant 2 ( <i>C</i> index: 0.528)	80.0	3
(3rd ranked participant in prediction accuracy)		
Neural network model (Participant 1)	78.3	7
Neural network model (Participant 13)	78.3	7
Tree structure model (Participant 2)	78.3	7
Participant 1 ( <i>C</i> index: 0.245)	78	10
(4th ranked participant in prediction accuracy)		

network model. However, the hit ratio in the linear regression model also shows a significant difference between the two groups. Therefore, we cannot accept Hypothesis 1, which states that valid non-linearity of decision-making is positively related to the predictive accuracy of a non-linear model, but not positively related to the predictive accuracy of a linear model. We can only infer that the valid non-linearity has an effect on predictive validity based upon the results of non-linear models. We may also assume that MPS is more suitable than the hit ratio because MPS is used for continuous value. If the above assumption is correct, and if we could give more weight to MPS than hit ratio, we might conjecture that no significant relationship exists between valid non-linearity and the predictive validity of the linear model.

#### 4.3. Comprehensive comparison

Table 5 shows the top ten predictive models and human participants based on hit ratio. Four participants and six models — three tree structure models and three neural network models — are included, but none of the linear regression models is included. It can be interpreted that as an expert's predictive accuracy increases, the modeling effect of linear models decreases. The highest hit ratio (83.3%) is achieved by the tree structure model for participant #3, and this value is even higher than actual human judgment for this participant. The predominance of a model over human judgment is also found for participant 1, whose neural network model (7th) outperforms human prediction (10th). *C* indexes of participants (#1, #2, #3, #9, and #13), whose models are ranked in the top 10, are the highest values among the 32 participants.

Table 6 lists the top ten predictive models and human participants based on MPS. The result is similar to Table 5. Three human experts (participants #2, #3, and #9) and seven non-linear models (four tree structure models and three neural network models) are included. However, none of the linear models is included. The highest predictive accuracy was achieved by the tree structure model of participant #3, the same result is shown in Table 5. The predominance of models over human experts was found in the case of participant #3, where his/her model (1st) outperforms his/her judgment (2nd), and in the case of participant #9, whose neural network model (6th) outperforms his/her judgment (8th). The same participants (#1, 2, 3, and 9) whose models are ranked top ten in Table 5 are also listed in Table 6. The results of Tables 5 and 6 show that there is valid non-linearity contributing to predictive accuracy in decision-making behavior and that non-linear models reflect non-linearity of behavior better than the linear model.

To shed light on this analysis, we measured the correlation between the predictive accuracy of all the participants ( $n=32$ ) and that of each decision model of the participants by hit ratio and  $R_a$ . We also measured the correlation between participants' predictive accuracy and the modeling effect of each decision model. Modeling effect is measured as the incremental difference between the model's predictive accuracy and the participant's predictive accuracy. Hence, it shows to what degree the accuracy of the model is greater than the accuracy of the human participant. Generally, it is believed that the higher the participant's predictive accuracy, the lower the modeling effect would be. Therefore, analyzing the correlation between a participant's predictive accuracy and modeling effect allows us to find out which modeling effect significantly decreases

Table 6

Top ten ranking of prediction accuracy using MPS

Model/participant	MPS	Rank
Tree structure model (Participant 3)	0.1387	1
Participant 3 ( <i>C</i> index: 0.534)	0.1430	2
(Top ranked participant in prediction accuracy)		
Neural network model (Participant 3)	0.1434	3
Tree structure model (Participant 2)	0.1548	4
Participant 2 ( <i>C</i> index: 0.528)	0.1548	4
(2nd ranked participant in prediction accuracy)		
Tree structure model (Participant 9)	0.1583	6
Neural network model (Participant 2)	0.1608	7
Neural network model (Participant 9)	0.1647	8
Participant 9 ( <i>C</i> index: 0.310)	0.1655	9
(3rd ranked participant in prediction accuracy)		
Tree structure model (Participant 1)	0.1672	10

Table 7

Correlation between participants' prediction accuracy and modeling effect

Model prediction accuracy and modeling effect		Participants' prediction accuracy	
		Hit ratio	$R_a$
Model prediction accuracy	Linear regression model	0.5866	0.5318
	Tree structure model	0.6005	0.5521
	Neural network model	0.8729	0.8006
Modeling effect	Linear regression model	−0.3925*	−0.3172*
	Tree structure model	−0.1063	−0.0557
	Neural network model	−0.1105	−0.0127

\* &lt;0.1.

as the participant's predictive accuracy increases. In this analysis, MPS is not used because it has a negative relationship to other prediction measurements. In addition, MPS is very highly correlated (−0.8598 in Table 2) with the hit ratio and  $R_a$ , which we do use in our analysis.

Table 7 shows the results of the analysis. The neural network model is most affected by the participant's predictive accuracy (correlation coefficient of 0.8729 with hit ratio and 0.8006 with  $R_a$ ). Generally, the modeling effect is negatively related to the participant's predictive accuracy. However, this negative correlation is found to be statistically significant only in the linear regression model (−0.3925 in hit ratio and −0.3172 in  $R_a$ ). This is consistent with the results found in Table 5 and Table 6, implying that the modeling effect of a linear model significantly deteriorates when the predictive validity of human judgment is high. We could interpret that this happens because the linear model can not reflect valid non-linearity, which ensures the predictive validity of the human judgment (or experts). This leads us to accept our second hypothesis: As the validity of a subject's decision-making increases, the modeling effect of a linear model significantly decreases to a greater extent than the modeling effect of a non-linear model.

## 5. Conclusion

It is believed that human decision-making behavior is quite non-linear and that using a non-linear model would be more effective in decision-making [12,17,19,24]. In this study, to test this hypothesis, we developed an expert decision-making model that shows how human experts use linear and/or non-linear models; and we investigated the relationship between valid non-linearity and predictive accuracy. We discovered several interesting results. First, we found a significant relationship between valid non-linearity and predictive accuracy through the ana-

lysis of human decision-making behavior using hit ratio, MPS, and  $C$  index of the lens model analysis. Second, non-linear models showed a higher predictive accuracy than linear models when valid non-linearity was inherent in human decision-making behavior. Third, through analyzing the modeling effects, we found that there is a significant negative relationship between the valid non-linearity of human decision-making behavior and the incremental predictive accuracy of linear models. We interpreted the negative correlation to mean that the more valid non-linearity there is in human judgment, the less effective the use of linear models would be. Hence, the negative correlation supports Libby's argument [20], that as expert validity increases, the advantage of using linear models decreases. The negative correlation could also present a possible explanation for the conflicting results concerning superiority of linear models in human judgment-modeling research. It seems that linear models outperformed non-expert judgment and non-linear models for non-experts for two reasons: There was significant noise in the input data and/or the human subjects used in the research were not domain experts. Consequently, it would be better to select the modeling algorithm based on the analysis of the characteristics of input data and/or the consideration of human expert decision-making behavior.

We believe the results of our research raises research issues in the field of expert systems and decision models. First, the lens model divides prediction accuracy into linear and non-linear components and provides different evaluation criteria. A lens model can answer the question of whether a non-linear model can show better predictive accuracy than a linear model or even human expert judgment.

The second issue involves knowledge management. Knowledge base is the core component of knowledge-based models. The fact that the predictive validity of experts, which provides a knowledge base, conveys a critical role in the predictive accuracy of the model has been proven by much research [19,21,37]. Using a lens model, we verified the predictive validity and valid non-linearity of experts. Based on our results, we believe that training experts could improve their predictive validity, which would eventually improve the predictive accuracy of the decision-making model.

Decision-making research, combined with expert system research, can expand to a variety of applications such as finance, medical science, and credit rating. However, more advanced decision model development requires systematic and synthesized study, which includes empirical studies of various areas and circumstantial factors affecting a system's predictive accuracy.

We hope that this research will be helpful to future research in considering additional circumstantial factors.

However, two limitations of this research should be noted. First, this research simplified the task of bankruptcy prediction by using only ten financial ratios. Such qualitative factors as management reputation and economic situation could not be included in this research design. In reality, such factors may play a critical role in predicting bankruptcy of a firm. Second, the evaluation of the case examples by the participants from the field (16 CPAs) was executed independently without our control. Since the task (the evaluation of 70 cases) was extremely time-consuming, it is possible that consistency during the whole evaluation process was not to be maintained. This limitation constrains the generalizability of the results.

### Acknowledgements

The authors thank Professor Andrew B. Whinston, Professor Jae Kyu Lee, Professor Ho Gun Lee and anonymous referees for their helpful comments on earlier drafts of the paper. All errors are those of the authors. This paper is supported by the Research Fund of the University of Seoul in 2006.

### References

- [1] E. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23 (4) (1968).
- [2] L. Beach, T. Mitchell, A contingency model for the selection of decision strategies, *Academy of Management Review* 3 (3) (1978).
- [3] A. Belkaoui, Lens studies in accounting, in: A. Belkaoui (Ed.), *Human Information Processing in Accounting*, Quorum Books, 1989.
- [4] R. Billings, S. Marcus, Measures of compensatory and non-compensatory models of decision behavior, *Organizational Behavior Human Performance* 31 (3) (1983).
- [5] H. Braun, J. Chandler, Predicting stock market behavior through rule induction, *Decision Sciences* 18 (3) (1987).
- [6] E. Brunswick, *The Conceptual Framework of Psychology*, University of Chicago Press, 1952.
- [7] C.J. Casey, Prior probability disclosure and loan officers' judgment: some evidence of the impact, *Journal of Accounting Research* 21 (1) (1983).
- [8] H.M. Chung, M.S. Silver, Rule-based expert systems and linear models: an empirical comparison of learning-by-examples methods, *Decision Sciences* 23 (3) (1992).
- [9] R. Cooksey, *Judgment Analysis: Theory, Methods, and Applications*, 1996 California, San Diego.
- [10] R. Dawes, B. Corrigan, Linear models in decision-making, *Psychological Bulletin* 81 (2) (1974).
- [11] H. Einhorn, The use of nonlinear, noncompensatory models in decision-making, *Psychological Bulletin* 77 (3) (1970).
- [12] H. Einhorn, Expert measurement and mechanical combination, *Organizational Behavior Human Performance* 20 (2) (1972).
- [13] H. Einhorn, D. Kleimuntz, B. Kleimuntz, Linear regression and process — tracing models of judgment, *Psychological Review* 86 (6) (1979).
- [14] D.H. Fisher, K.B. McKusick, An empirical comparison of ID3 and back-propagation, Technical Report CS-88-14, Dept. of Computer Science, Vanderbilt University, 1989.
- [15] K. Harmond, C. Hursch, F. Todd, Analyzing the components of clinical inference, *Psychological Review* 71 (6) (1964).
- [16] C. Harris, An Expert Decision Support Systems for Auditor Going Concern Evaluation, PhD Dissertation, University of Texas at Arlington (1989).
- [17] C.N. Kim, R. McLeod Jr., Expert, linear models, and nonlinear models of expert decision-making in bankruptcy prediction, *Journal of Management Information Systems* 16 (1) (1999).
- [18] C.N. Kim, M. Chung, D. Paradise, Inductive learning of expert decision-making: a decision strategy perspective, *Decision Support Systems* 21 (2) (1997).
- [19] K. Levi, Expert systems should be more accurate than human experts, *IEEE Transactions on Systems, Man, and Cybernetics* 19 (3) (1989).
- [20] R. Libby, Man versus model of man: some conflicting evidence, *Organizational Behavior Human Performance* 16 (1) (1976).
- [21] R. Libby, *Accounting and Human Information Processing: Theory and Applications*, 1981 New Jersey, Englewood.
- [22] W. Messier, J. Hansen, Inducing rules for expert system development: an example using default and bankruptcy data, *Management Science* 34 (12) (1988).
- [23] M. O'Connor, W. Remus, K. Lim, Improving judgmental forecasts with judgmental bootstrapping and task feedback support, *Journal of Behavioral Decision Making* 18 (2005).
- [24] R.W. Olshavsky, Task complexity and contingent processing in decision-making: a replication and extension, *Organizational Behavior Human Performance* 24 (3) (1979).
- [25] E. Patuwo, M. Hu, M. Hung, Two-group classification using neural networks, *Decision Sciences* 24 (4) (1993).
- [26] J. Quinlan, Discovering rules by induction from large collection of examples, in: D. Michie (Ed.), *Expert Systems in Microelectronic Age*, 1979, England, Edinburgh.
- [27] A. Schepanski, Tests of theories of information processing behavior in credit judgment, *Accounting Review* 58 (3) (1983).
- [28] T. Stewart, Judgmental analysis: procedures, in: B. Brehmer, R.B. Joyce (Eds.), *Human Judgment: The SJT View*, North-Holland, Amsterdam, 1988.
- [29] T. Stewart, Improving reliability of judgmental forecasts, in: J.S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, 2001, Massachusetts, Boston.
- [30] O. Svenson, Process descriptions of decision-making, *Organizational Behavior Human Performance* 23 (1) (1979).
- [31] K. Tam, M. Kiang, Managerial applications of neural networks: the case of bank failure predictions, *Management Science* 38 (7) (1992).
- [32] L.R. Tucker, A suggested alternative formulation in the development of Hursch, Hammond, and Hursch, and by Hammond, Hursch, and Todd, *Psychological Review* 71 (6) (1964).
- [33] A. Velido, P. Lisboa, J. Vaughan, Neural networks in business: a survey of applications (1992–1998), *Expert Systems with Applications* 17 (1999).
- [34] B. Wong, V. Lai, J. Sam, A bibliography of neural network business applications research: 1994–1998, *Computers & Operations Research* 27 (2000).
- [35] J.F. Yates, External correspondence: decompositions of the mean probability score, *Organizational Behavior Human Performance* 30 (1982).

- [36] A. Zacharakis, G. Meyer, A lack of insight: do venture capitalists really understand their own decision process, *Journal of Business Venturing* 13 (1998).
- [37] I. Zimmer, A lens study of the prediction of corporate failure by bank loan officers, *Journal of Accounting Research* 18 (2) (1980).

**Choong Nyoun Kim** (cnkim27@uos.ac.kr) is a professor in the Management Department at the University of Seoul. He received his M.B.A. from the University of Missouri at Columbia and his PhD from Texas A&M University. His research interests include decision support systems and expert systems, human information processing, and machine learning. His research findings have been published in *Decision Support Systems*, the *Journal of Global Information Management*, and the *Journal of Management Information Systems*.

**Kyung Hoon Yang** (yang.kyun@uwlax.edu) is currently an assistant professor in the Department of Information Systems at the University of Wisconsin–La Crosse. He received his PhD in management information systems from Purdue University and was a professor at Chungang University in Korea. His major research interests include knowledge transfer and knowledge-based organization. He has published more than 15 articles in journals, which include *Expert Systems with Applications* and *Omega*.

**Jaekyung Kim** (jkim6@unl.edu) is a PhD candidate in Information Systems at the University of Nebraska–Lincoln. He received his M.B.A. from Miami University of Ohio. His research focuses on decision support systems, electronic commerce, and knowledge management. His work has been published in journals such as *Omega*, *Expert Systems with Applications* and the *International Journal of Knowledge Management*.