An Applied Study Using the Analytic Hierarchy Process to Translate Common Verbal Phrases to Numerical Probabilities

MADJID TAVANA[1] and DENNIS T. KENNEDY[1]
[1]*La Salle University, USA*

BARBARA MOHEBBI[2]
[2]*Fairleigh Dickinson University, USA*

ABSTRACT

This is an applied study of how to develop a standardized set of useful verbal probability phrases for communication purposes within an expert community. The analysis extends the previous research in two ways. First, the Analytic Hierarchy Process (AHP) is used to assess the relative weights associated with the verbal phrases employed by a group of thirty financial strategy experts at a major Wall Street firm. Second, a quadratic least-squares technique is used to map these relative weights onto a subjective probability scale. The result was a consistent scaling of probabilistic phrases that the analysts prefer and actually use. This methodology can be used to minimize the problems associated with the use of probabilistic phrases by a group of experts who interact daily and who share assumptions, working knowledge and values. © 1997 by John Wiley & Sons, Ltd.

INTRODUCTION

Subjective probabilities play an important role in human decision making. Budescu and Wallsten (1987), Wallsten and Budescu (1990) and Hamm (1991) assert that many people claim to use nonnumerical phrases to express subjective probabilities because it is how they think and because nonnumerical phrases better express their perceived uncertainty than numbers. Studies by Wallsten *et al.* (1993b), Erev and Cohen (1990), and Brun and Teigen (1988) confirm a preference to receive numerical expressions of uncertainty but to convey probability judgments verbally in a decision-making task.

However, the empirical research indicates little interpersonal agreement on the meaning of probabilistic phrases. Brun and Teigen (1988) and Budescu and Wallsten (1990) provide excellent reviews and evaluations of the empirical research on the use of probabilistic phrases. The previous

research, together with the results of their studies, led them to conclude that there is a large variability in numerical estimates associated with verbal phrases. However, Budescu and Wallsten (1985, 1990) caution against the strong conclusion that the use of nonnumerical phrases increases confusion while the use of numbers enhances communication. It is possible that different results would occur with a group of experts who interact daily and who share assumptions, working knowledge and values.

Beyth-Marom (1982) explored this possibility with an experiment that took place in a professional forecasting organization with experts who were accustomed to giving verbal probability assessments. Although individual subjects were relatively consistent in their assignment of numerical probabilities to nonnumerical probabilistic phrases, there was considerable disagreement among the experts in the interpretation of most of the verbal probability statements. The variability was even higher when the experts were asked to evaluate the verbal probability expressions in the context of publications by their organization. Another interesting result was that not all phrases were equally vague.

Budescu and Wallsten (1990) examined the use of probability phrases by forecasters communicating uncertainty to decision makers. The subjects were university students who used numerical and verbal probabilities to describe the probability that a dart would land on a shaded section of a circle. The decision makers made bids based on the numerical or verbal description. The results were consistent with the prior research findings of individual-subject studies (Budescu *et al.*, 1988; Rapoport *et al.*, 1987; Wallsten *et al.*, 1986). The responses to verbal descriptors were more variable than the responses to numerical descriptors. In addition, the results support the findings of Fillenbaum *et al.* (in press) that decision makers interpret probability phrases to have wider range of meaning than the forecaster intended.

These results suggest organizations should use numerical expressions of probability rather than verbal ones. However, the resistance to such a change can be enormous. A compromise is a scale of verbal probabilistic phrases with explicit numerical translation. After observing the inconsistent way probability expressions are used in the field of accounting and auditing, Chesley and Wier (1985, p. 41) conclude, 'The good news is that the problem has a solution: the assignment of a suggested range of probabilities to standard uncertainty expressions'. Similarly, Olsen and O'Neill (1988, p. 72) recommend that financial analysts provide their clients with a formal scale indicating a range of numerical probabilities associated with the probabilistic phrases being used. Budescu and Wallsten (1990, p. 259) suggest that communications in actual decision situations will not suffer so long as the decision maker treats the selected phrase as intended by the forecaster.

In a review of 20 different studies, Mosteller and Youtz (1990) noted only modest variation in the average numerical values associated with most probabilistic expressions and stability of meanings over 20 years. They suggest that the next step is 'to offer codifications and see how satisfactory people find them' (p. 2). A number of commentators agree that quantification of uncertainty expressions would be useful (Tanur, 1990; Kadane, 1990) especially in those fields where they are used to describe scenarios and forecasts (Wolf, 1990). Kadane (1990) proposed a scale based on the Mosteller and Youtz (1990) data and concluded with the challenge 'Now it's your turn' (p. 20). However, meaning occurs from usage in a context, so any attempt at associating verbal and numerical probabilities should be done in context (Clark, 1990; Fox, 1986; Tanur, 1990). In contrast to the perspective of Erev *et al.* (1991) that vagueness and ambiguity are essential in certain situations to facilitate heterogeneous choices, Cliff (1990) suggested that the context most conducive to standardizing terminology is characterized by isolation of the communicators, specificity of the referent and social or economic penalty for linguistic error.

Recent studies have proposed methods for obtaining and numerically evaluating verbal probabilistic phrases. Hamm (1991) proposed a method that avoids the problems of an indefinitely large vocabulary and individual disagreements about their meaning. The method required the subjects to use phrases from a list that spans the full range of probabilities. The list was constructed by reviewing studies

that elicited numerical values for verbal probabilistic expressions (Budescu and Wallsten, 1985; Lichtenstein and Newman, 1967; Simpson, 1944; Shanteau, 1974; Wallsten *et al.*, 1986). Introductory psychology students used the list to describe a problem scenario. Then, the subjects assigned numerical values to the phrases. The solution was produced by applying Bayes' Theorem (Hamm, 1988). The results showed that the use of an ordered list of verbal probabilistic phrases enabled subjects to select phrases that were more accurate answers for the word problems. The results were viewed as supporting the recommendations of Kong *et al.* (1986) to improve the use of verbal probabilities by measuring what people mean by phrases, publishing these meanings and training people to use the terms with these accepted meanings. However, the process of translating phrases was seen as a burden for the subject and the researcher.

Wallsten *et al.* (1993a) conducted an experiment in which university students were asked to select and scale their own set of verbal expressions. The subjects used pair-comparisons to assign numerical values to the verbal phrases they preferred to use. The results revealed little consensus between the subjects in their selection of verbal phrases. Therefore, the analysis was 'done at the individual subject level' (p. 182) and indicated that an individual's responses using numerical and verbal probabilities were virtually identical. The results suggest that verbal forecasting requires analysts select and scale their own vocabulary and that it be known to the users of the forecast. While the pair-comparison procedure gave a more complete representation of an individual's understanding of a verbal probability expression, it was described as arduous.

This study accepts Kadane's (1990) challenge and explores the development of a probabilistic phrases scale proposed by Mosteller and Youtz (1990). The analysis extends the previous research in two ways. First, the Analytic Hierarchy Process (AHP) is used to assess the relative weights of the verbal phrases employed by a group of thirty experts in financial strategy at a major Wall Street firm. Second, a quadratic least-squares technique is used to map these relative weights onto a subjective probability scale. The analysis is conducted in three phases. In the initial phase, the experts are asked to consider eleven groups of phrases with similar numerical values and to select the phrase which is most representative of each group. Next, the experts are asked to rank the verbal expressions according to their subjective probability. Then, AHP is used to evaluate the relative weights of the probabilistic phrases. Finally, a quadratic least-squares model is used to translate these relative weights into a numerical probability scale.

Probabilistic phrases and other communicative functions

Other investigators have observed that verbal uncertainty expressions are used for semantic and communicative functions besides expressing numerical probabilities. For example, Brun and Teigen (1988) identified an affective intensity dimension and concluded, 'It is likely that words differ in other respects as well and perform a communication function that canmnot be fully replaced by "neutral" numbers'. Subsequently, Teigen and Brun (1995) demonstrated that verbal phrases exhibit a directional aspect. They found that some terms, such as possible and probable function as affirmative phrases emphasizing the occurrence of an event. In contrast, other terms, such as unlikely and doubtful serve as negative phrases suggesting non-occurrence. They noted that similar phenomena have been described in psycholinguistic studies by Champaud and Bassano (1987), Moxey and Sanford (1987), and Moxey *et al.* (1990). Fox (1986) suggested that there are different types of uncertainty including possibility, plausibility, and probability. He observed that verbal terms are used to express various degrees of uncertainty, but they 'do not make the underlying uncertainty semantics explicit' (p. 448). He emphasized that there may be no universally correct calculus of uncertainty, but knowledge about uncertainty should be made explicit. In contrast, Erev *et al.* (1991) suggest that precise communications

are not always desirable and that vagueness and ambiguity are essential in certain situations to facilitate heterogeneous choices.

The context within FSG exhibits the characteristics identified by Cliff (1990) as most conducive to standardizing language and described by Wolf (1990) as where it would be most useful rather than the characteristics where vagueness and ambiguity are essential (Erev *et al.*, 1991). Therefore, the study began with an examination of FSG internal communications. Their length ranged from two to six pages and they contained a variety of probabilistic phrases. Further analysis revealed that the analysts assigned the phrases a wide range of overlapping probabilities. Based on these initial findings, the study proceeded with the application of a method to help the analysts identify and scale a set of verbal probability expressions that met their needs. The researchers and FSG management were not concerned with other communicative functions because the entire report was available for their expression. This represents a limitation of the current study and an opportunity for future research in an applied setting.

METHOD

Subjects
The participants were thirty expert financial strategists at the Financial Strategies Group (FSG) of a major Wall Street firm. FSG was established to provide state-of-the-art analytical tools to the firm's clients, as well as to the firm's sales, trading and banking areas. The group members interact frequently during a workday, and their communication frequently includes the use of nonnumerical phrases to express perceived uncertainty.

The FSG includes five different groups of expert financial strategists each with a director overseeing the operations of the group. Experts working in each group develop business forecasts for their product areas and these forecasts are folded into an overall forecast. The sample included 12 females and 18 males. The ages range from 21 to 46 with an average of 26 years. The level of education varies from doctorates among the group managers, to masters and bachelor degrees among the forecasters and analysts. The degrees represent a number of disciplines including statistics, economics, finance, computer science, and mathematics. Tenure in the group ranges from three months to nine years. With the diversity in gender, age, education, tenure and role in FSG there is the potential for large variability in numerical estimates associated with nonnumerical probabilistic phrases.

Procedure
The data for this study were collected with the three questionnaires presented in Appendices A–C. The instruments were administered on three separate occasions. Prior to the administration of each questionnaire, the researchers met with the participants to discuss the purpose of the instrument. The first questionnaire included a comprehensive list of the probabilistic and frequentistic expressions used in prior research conducted by Brun and Teigen (1988), Budescu and Wallsten (1985), and Lichtenstein and Newman (1967). The 76 probabilistic phrases were divided into eleven groups according to the similarity of the numerical values assigned in previous studies. The instructions asked the experts to select one verbal expression whose meaning was most representative of the group. During the discussion with the participants, they were asked to select the phrase in each group that most clearly captured the probability conveyed by the phrases in that group. To minimize the possibility of bias, the groups of phrases were randomly placed on the questionnaire and the phrases were randomly distributed within each group.

Two weeks later, a second questionnaire was distributed. It contained the most frequently selected probabilistic phrase from each of the eleven groups on the first questionnaire. However, it included no frequentistic phrase because none were selected by the analysts most often as representative of a group.

Exhibit 1. AHP pairwise comparison scale

AHP verbal scale	AHP numerical scale
Extremely more probable	9
Very strongly more probable	7
Strongly more probable	5
Moderately more probable	3
Equally probable	1
Moderately less probable	1/3
Strongly less probable	1/5
Very strongly less probable	1/7
Extremely less probable	1/9

The purpose of this questionnaire was to identify an ordinal ranking associated with these phrases. The experts were asked to rank these probabilistic expressions from 1 to 11 with 1 assigned to the expression indicating the highest probability and 11 to the expression indicating the lowest probability. As before, these representative phrases were placed randomly on the questionnaire.

A third questionnaire was used to capture the experts' judgments about the relative likelihood implied by the nonnumerical probabilistic phrases. This questionnaire is based on AHP. AHP was introduced by Saaty (1972, 1977a) to assist a decision maker in evaluating complex judgmental problems. AHP helps the decision maker assign numerical values to qualitative attributes by making trade-offs among them. The process is confined to a series of pairwise comparisons using the scale presented in Exhibit 1. Saaty (1972) argues that a decision maker naturally finds it easier to compare two things than to compare all the items in a list. AHP also evaluates the consistency of the decision maker and allows for the revision of the responses. Because of the intuitive nature of the process and its power in resolving the complexity in a judgmental problem, AHP has been applied to many diverse decisions. A comprehensive list of the major applications of AHP, along with a description of the method and its axioms, can be found in Saaty (1972, 1977a, b, 1980, 1990), Weiss and Rao (1987) and Zahedi (1986). A mathematical summary of AHP is presented in Appendix D.

In order to reduce the number of pairwise comparisons, the third questionnaire was divided into two parts. Phrases identified with a 'toss-up' or higher probability were placed on one questionnaire while phrases identified with a 'toss-up' or lower probability were placed on a second questionnaire. Part 1 contained a series of pairwise comparisons between the phrases identified in the second questionnaire with a perceived numerical probability of 50% or less. Part 2 contained another series of pairwise comparisons between the phrases identified with a perceived numerical probability of 50% or more. Within each part of the questionnaire, the pairwise comparisons were listed randomly.

RESULTS

A preliminary test was conducted to motivate FSG management to authorize the project. This initial phase examined: (1) the occurrence of verbal and numerical probability expressions in FSG internal communications and (2) the variability in the analysts' interpretations of the verbal phrases. The researchers were permitted to examine 12 months of internal communications at FSG. Typically, the communications included several verbal probability phrases but rarely included numerical probabilities. In consultation with FSG management, the researchers selected ten representative messages that include 16 different verbal phrases. To maintain confidentiality, the complete text of any FSG communication cannot be disclosed; however, Appendix E contains some relevant excerpts.

Exhibit 2. Descriptive statistics for the preliminary test ($n = 30$)

Verbal phrase	Mean	Std dev.	Min.	Max.	Range
Probably	74.66	9.72	54	85	31
Very probable	82.50	8.83	63	94	31
Not very probable	25.83	6.73	10	33	23
Very doubtful	22.83	6.47	12	44	32
Good chance	82.18	10.70	60	95	35
Very unlikely	28.44	11.25	0	44	44
Fair chance	32.26	15.86	18	66	48
Unlikely	31.42	10.41	5	45	40
Small probability	21.76	7.88	5	33	28
Often	77.17	9.43	63	99	36
Predictable	47.95	17.56	28	77	49
Usually	73.71	10.35	54	95	41
Frequently	81.16	11.31	54	99	45
Rare	13.81	8.97	0	42	42
Possibly	23.81	8.96	10	50	40
Always	83.12	8.78	66	99	33

Next, the 30 FSG analysts estimated the numerical probability that each phrase implies within the context of the communication. Exhibit 2 shows that the analysts assigned a wide range of probabilities to the phrases. The range for two phrases is less than 29%, for six phrases the range is between 30% and 39%, and for the remaining eight phrases the range is over 40%. In addition, the ranges overlapped considerably. These results convinced FSG management to proceed with the project.

Based on the results of prior research, Questionnaire 1 (Appendix A) presented a list of 76 commonly used verbal probabilistic phrases divided into 11 groups with similar perceived probabilities. Exhibit 3 presents a frequency distribution indicating the number of participants who designated each of the verbal expressions as the most representative phrase of its group. The results indicate that the 11 most frequently selected phrases were: 'possible', 'small possibility', 'somewhat likely', 'somewhat doubtful', 'likely', 'impossible', 'quite certain', 'small chance', 'certain', 'toss-up', and 'very likely'.

Next, Questionnaire 2 (Appendix B) was distributed. With this questionnaire, the experts were asked to rank eleven verbal phrases identified in the first questionnaire according to their perceived probabilities. Exhibit 4 presents the statistical results. 'Certain' was ranked as the phrase indicating the highest probability while 'impossible' was identified as the phrase indicating the lowest probability. 'Toss-up' was ranked in the middle. 'Certain', 'quite certain', 'very likely', 'likely', and 'somewhat likely' were associated with a probability greater than 'toss-up'. In contrast, 'possible', 'somewhat doubtful', 'small chance', 'small possibility' and 'impossible' were considered to represent a probability less than 'toss-up'.

Questionnaire 3 (Appendix C) presented the experts with a series of fifteen pairwise comparisons between the phrases with a 'toss-up' or lower probability in Part 1 and another series of fifteen pairwise comparisons between the phrases with a 'toss-up' or higher probability in Part 2. The responses were entered into a computer program that performed the AHP computations to derive the relative weights associated with each of the phrases and a consistency ratio for each of the experts (Saaty, 1977a, b, 1980).

The statistical results for Part 1 are presented in Exhibit 5 and for Part 2 in Exhibit 6. The inconsistency ratio for each series of pairwise comparisons indicates very little within-subject variability (Saaty, 1980, 1990). This implies that the individual experts were relatively consistent in their assessment of the relative weights identified by the phrases. Furthermore, the small standard deviations

Exhibit 3. Response frequency for verbal phrases in Questionnaire 1 ($n = 30$)

Verbal phrase (Frequency)	
Uncertain (4)	Perhaps (3)
Inconclusive (0)	Not Certain (3)
A Certain Hope (0)	**Possible (12)**
Possibly (5)	Unpredictable (3)
Improbable (2)	Rare (0)
Not Likely (3)	**Small Possibility (10)**
Barely Possible (1)	Very Doubtful (3)
Highly Improbable (1)	Not Probable (3)
Small Probability (4)	Rarely (1)
Quite Unlikely (0)	Very Unlikely (2)
Rather (0)	Slightly More Than Half the Time (0)
Better Than Even (3)	Predictable (4)
Somewhat Likely (18)	Slight Odds in Favour (5)
Somewhat Unlikely (11)	**Somewhat Doubtful (19)**
Probable (4)	Frequently (2)
Pretty Good Chance (1)	Good Hope (0)
Fairly Likely (2)	Probably (5)
Common (0)	Rather Likely (1)
Likely (11)	Often (2)
Good Chance (2)	Predictable (0)
No Chance (5)	Never (4)
Not Possible (6)	**Impossible (15)**
Quite Likely (8)	Usually (0)
Most Possibly (5)	Great Chance (5)
Very Good Hope (0)	**Quite Certain (12)**
Small Chance (11)	Improbable (3)
Small Likelihood (0)	Not Very Probable (2)
Seldom (0)	Rather Unlikely (4)
Fairly Unlikely (2)	Quite Doubtful (0)
Uncommon (0)	Unlikely (7)
Doubtful (1)	Usually Not (0)
Not Much Chance (0)	A Small Hope (0)
Always (9)	**Certain (21)**
Fair Chance (6)	**Toss-up (16)**
Small Doubt (5)	Slight Odds Against (3)
Fighting Chance (0)	Danger (Risk) (0)
Slightly Less Than Half the Time (0)	
Very Likely (18)	Very Probable (7)
Highly Probable (5)	

and narrow ranges reveal a small degree of between-subject variability. There appears to be considerable interpersonal agreement on the meaning of the phrases. In addition, the means and standard deviations indicate little overlap in the use of the terms. This suggests that the participants assigned distinct meaning to the phrases in Questionnaire 3.

Exhibit 4. Mean ranks of selected verbal phrases for Questionnaire 2 ($n = 30$)

Verbal phrase	Mean	Std dev.	Min.	Max.
Certain	1.07	0.254	1	2
Quite Certain	2.10	0.403	1	3
Very Likely	2.83	0.461	1	3
Likely	4.10	0.305	4	5
Somewhat Likely	4.90	0.305	4	5
Toss-up	6.23	0.626	6	8
Possible	7.13	0.507	6	9
Somewhat Doubtful	7.70	0.702	6	9
Small Chance	9.17	0.531	8	10
Small Possibility	9.77	0.430	9	10
Impossible	11.00	0.000	11	11

Exhibit 5. Descriptive statistics for Questionnaire 3 — Part 1 ($n = 30$)

Verbal phrase	Mean relative weight	Std dev.	Min.	Max.
Toss-up	0.341	0.017	0.299	0.366
Possible	0.241	0.020	0.201	0.283
Somewhat Doubtful	0.168	0.015	0.131	0.195
Small Chance	0.129	0.013	0.106	0.157
Small Possibility	0.080	0.007	0.065	0.090
Impossible	0.041	0.004	0.035	0.049
Inconsistency	**0.065**	**0.018**	**0.028**	**0.095**

Exhibit 6. Descriptive statistics for Questionnaire 3 — Part 2 ($n = 30$)

Verbal phrase	Mean relative weight	Std dev.	Min.	Max.
Certain	0.346	0.015	0.320	0.378
Quite Certain	0.239	0.015	0.201	0.270
Very Likely	0.175	0.015	0.141	0.205
Likely	0.122	0.015	0.094	0.167
Somewhat Likely	0.077	0.008	0.064	0.091
Toss-up	0.047	0.003	0.036	0.049
Inconsistency	**0.063**	**0.018**	**0.019**	**0.092**

The next step was to map the relative weight associated with each of the phrases into the numerical probability scale. Assuming that the probabilities are a quadratic function of the weights and they are monotonically increasing function with respect to weights, two models were constructed similar to the model developed by Tobler (1979). These models use a quadratic curve and least-square on spacing to convert the relative weights in Exhibits 5 and 6 into numerical probabilities. In the first model, the phrases 'impossible' and 'toss-up' were used as anchors representing the 0% and 50% probabilities. Given the relative weight of 0.041 for 'impossible' and 0.341 for 'toss-up', the model was used to calculate the probability for each of the other phrases associated with a relative weight between 'impossible' and 'toss-up'. Exhibit 7 shows the calculated probabilities for each phrase and a suggested probability for use in the development of a formal scale. For example, the relative weight of 0.241 for

Exhibit 7. Quadratic curve results for Questionnaire 3 — Part 1 ($n = 30$)

Verbal phrase	Mean relative weight	Calculated probability	Suggested probability	Change
Toss-up	0.341	50.00[a]	50	—
Possible	0.241	40.12	40	−0.12
Somewhat Doubtful	0.168	28.62	30	+1.38
Small Chance	0.129	21.00	20	−1.00
Small Possibility	0.080	9.95	10	+0.05
Impossible	0.041	0.00[a]	00	—

[a]Assumed values for anchor phrases to be used in Quadratic Curve and Least Square Spacing.

Exhibit 8. Quadratic curve results for Questionnaire 3 — Part 2 ($n = 30$)

Verbal phrase	Mean relative weight	Calculated probability	Suggested probability	Change
Certain	0.346	100.00[a]	100	—
Quite Certain	0.239	90.03	90	−0.03
Very Likely	0.175	80.16	80	−0.16
Likely	0.122	69.76	70	+0.24
Somewhat Likely	0.077	59.36	60	+0.64
Toss-up	0.047	50.00[a]	50	—

[a]Assumed values for anchor phrases to be used in Quadratic Curve and Least Square Spacing.

Exhibit 9. Verbal probabilistic expressions and perceived probability estimates

Verbal expression	Probability
Impossible	0.00
Small Possibility	0.10
Small Chance	0.20
Somewhat Doubtful	0.30
Possible	0.40
Toss-up	0.50
Somewhat Likely	0.60
Likely	0.70
Very Likely	0.80
Quite Certain	0.90
Certain	1.00

'possible' is converted into a 40.12% probability which would be adjusted to 40%. This represents an adjustment of only −0.12. Similar results were obtained for the other phrases.

Exhibit 8 presents the quadratic curve results for the phrases in Questionnaire 3 (Part 2). For this model, the terms 'toss-up' and 'certain' were the anchors representing the 50% and 100% probabilities. Given the relative weight of 0.047 for 'toss-up' and 0.346 for 'certain', the model estimated the probability for each of the other phrases associated with a relative weight between 'toss-up' and 'certain'. For example, the relative weight of 0.239 for 'quite certain' is converted into a probability of 90.03%. This calculated probability would be adjusted to 90% for use in the development of a formal scale. This represents an adjustment of only −0.03. Similar results were obtained for the other phrases. Exhibit 9 represents a complete listing of verbal probabilistic

Exhibit 10. Frequency of numerical and verbal probabilities

	Frequency	
	Before scale introduction	After scale introduction
Total numerical	3	0
Verbal expression:		
Small Possibility	0	5
Very Doubtful	0	1
Very Unlikely	2	0
Not Likely	4	0
Small Chance	1	5
Unlikely	4	0
Possible	4	6
Uncertain	2	1
Likely	6	8
Probable	5	0
Very Likely	7	7
Very Probable	2	0
Usually	3	0
Certain	1	2
Total verbal	41	35

expressions along with the numerical probabilities suggested for the development of a verbal probability scale.

The FSG management reviewed the results and decided to introduce the verbal probability scale to the analysts. After a discussion of the preliminary test results and the scale development, the analysts decided that they would use either numerical probabilities or the verbal phrases in the suggested scale. Subsequently, the researchers examined the FSG internal communications for 6 months to assess how the scale was actually being used by the analysts.

A review of these communications revealed that the incidence of numerical probabilities was extremely rare after the introduction of the scale. Furthermore, the variety of probabilistic phrases was reduced significantly. In a few instances, the analysts expressed their ignorance or feelings of doubt with phrases that were not on the scale. However, whenever they were expressing the probability of events or circumstances, the analysts used only phrases from the scale. Exhibit 10 presents data comparing the frequency of numerical and verbal probabilities in a randomly selected subset of 30 FSG communications before and after the introduction of the scale.

To examine their preferences, the analysts were asked to compare the verbal probability scale they had developed (alternative V) to using numerical probabilities only (alternative N) and to the prior circumstances that permitted verbal phrases without specific numerical translations but with the option of numerical probabilities (alternative P). The analysts were asked to express their preferences for these alternatives when receiving and sending communications. The comparisons were presented to the analysts as the hierarchy in Exhibit 11. The analysts used the Expert Choice software to make all the necessary pairwise comparisons three times during a nine-month period.

The results of the three AHP preference assessments are presented in Exhibit 11. The data within each element of the hierarchy are the geometric means of the analysts preferences from the three assessments and indicate that the analysts' preferences were quite stable during the time period. Overall, the FSG analysts expressed strong preferences for alternative V compared to either alternative P or alternative N, but they were indifferent between P and N. However, when contrasting receiving
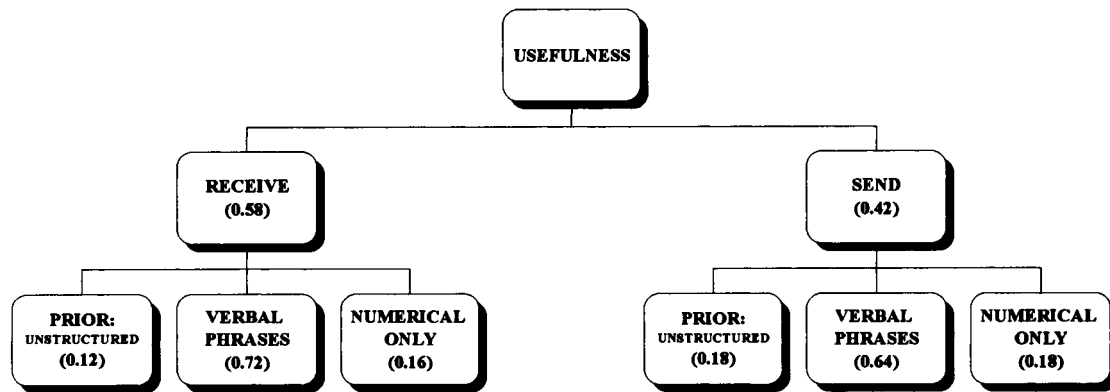
Exhibit 11. Preferences hierarchy and the results of three AHP assessments

and sending, the analysts seemed to indicate that the contribution of V was more important in receiving (0.58×0.72) when compared to its contribution to sending (0.42×0.64).

CONCLUSION

Research has confirmed that people prefer to use verbal phrases rather than numerical probabilities when conveying uncertainty but prefer to receive it numerically (Wallsten *et al.*, 1993b; Erev and Cohen, 1990; Fillenbaum *et al.*, in press; Budescu and Wallsten, 1990; Rapoport *et al.*, 1987). In addition, they attach different probabilities to individual phrases, and they overlap the meaning between them. Recent studies (Wallsten *et al.*, 1993a; Hamm, 1991) have proposed methods for reducing the problems of verbal probability expressions. The procedures required individuals to select and scale the phrases they prefer to use. These studies revealed that scaling verbal expressions within a particular decision context improved the usefulness of the phrases in that setting. However, the procedures were described as arduous (Wallsten *et al.*, 1993a, p. 179) and a burden (Hamm, 1991, pp. 206 and 209).

The proposed method makes it easier to codify the meaning that individuals assign to verbal probability expressions, to publicize these meanings and to train people to use the terms with these common meanings. The results demonstrate that professional colleagues are able to agree on the interpretation of probabilistic phrases with little overlap when they select a representative set of phrases and make comparisons among them in a systematic manner. Furthermore, assessments before and after the implementation of the verbal probability scale indicate that the verbal phrase scale is working at FSG. Specifically, the financial analysts restricted their verbal expressions of probability to the phrases in the scale. In contrast, the development and implementation of the verbal probability scale did not appear to constrain the analysts' use of uncertainty phrases for other communicative purposes. Finally, the financial analysts expressed a preference for the use of the verbal probability scale to either the use of numerical probabilities alone or to the unrestricted use of undefined verbal phrases that existed prior to this study.

ACKNOWLEDGEMENTS

REFERENCES

Beyth-Marom, R., 'How probable is probable? A numerical translation of verbal probability expressions', *Journal of Forecasting*, **1** (1982), 257–69.

Brun, W. and Teigen, K. H., 'Verbal probabilities: ambiguous, context-dependent, or both?' *Organizational Behavior and Human Decision Processes*, **41** (1988), 390–404.

Budescu, D. V. and Wallsten, T. S., Consistency in interpretation of probabilistic phrases', *Organizational Behavior and Human Decision Processes*, **36** (1985), 391–405.

Budescu, D. V. and Wallsten, T. S., 'Subjective estimation of precise and vague uncertainties', in Wright, G. and Ayton, P. (eds), *Judgmental Forecasting* (pp. 63–81), New York: Wiley, 1987.

Budescu, D. V. and Wallsten, T. S., 'Dyadic decisions with numerical and verbal probabilities', *Organizational Behavior and Human Decision Processes*, **46** (1990), 240–63.

Budescu, D. V., Weinberg, S. and Wallsten, T. S., 'Decisions based on numerically and verbally expressed uncertainties', *Journal of Experimental Psychology: Human Perception and Performance*, **14** (1988), 281–94.

Champaud, C. and Bassano, D., 'Argumentative and informative functions of French intensity modifiers "presque" (almost), "a peine" (just, barely) and "a peu pres" (about): An experimental study of children and adults', *European Bulletin of Cognitive Psychology*, **7** (1987), 605–31.

Chesley, G. R. and Wier, H. A., 'The challenge of contingencies: adding precision to probability', *CAMAGAZINE*, April (1985), 38–41.

Clark, H. H., 'Comment on "Quantifying probabilistic expressions" ', *Statistical Science*, **5** (1990), 12–16.

Cliff, N., 'Comment on "Quantifying probabilistic expressions" ', *Statistical Science*, **5** (1990), 16–18.

Erev, I. and Cohen, B. L., 'Verbal versus numerical probabilities: efficiency, biases, and the preference paradox', *Organizational Behavior and Human Decision Processes*, **45** (1990), 1–18.

Erev, I., Wallsten, T. S. and Neal, M. M., 'Vagueness, ambiguity, and the cost of mutual understanding', *Psychological Science*, **2** (1991), 321–4.

Fillenbaum, S., Wallsten, T. S., Cohen, B. and Cox, J. A., 'Some effects of available vocabulary and communication task on the understanding and use of non-numerical probability expressions', *American Journal of Psychology*, in press.

Fox, J., 'Three arguments for extending the framework of probability', in Kanal, L. N. and Lemmer, J. F. (eds), *Uncertainty in Artificial Intelligence*, Amsterdam: North-Holland, 1986.

Hamm, R. M., *Accuracy of Probabilistic Inference Using Verbal vs. Numerical Probabilities*, Chicago: Psychonomics Society Meetings (Conference) (1988).

Hamm, R. M., 'Selection of verbal probabilities: a solution for some problems of verbal probability expression', *Organizational Behavior and Human Decision Processes*, **48** (1991), 193–223.

Kadane, J., 'Codifying chance: Comment on "Quantifying Probabilistic Expressions" ', *Statistical Science*, **5** (1990), 2–30.

Kong, A., Barnett, G. O., Mosteller, F. and Youtz, C., 'How medical professionals evaluate expressions of probability', *New England Journal of Medicine*, **315** (1986), 740–5.

Lichtenstein, S. and Newman, J. R., 'Empirical scaling of common verbal phrases associated with numerical probabilities', *Psychonomic Science*, **9** (1967), 563–4.

Mosteller, F. and Youtz, C., 'Quantifying probabilistic expressions', *Statistical Science*, **5** (1990), 2–30.

Moxey, L. M. and Sanford, A. J., 'Quantifiers and Focus', *Journal of Semantics*, **5** (1987), 189–206.

Moxey, L. M., Sanford, A. J. and Barton, S. B., 'Control of attentional focus by quantifiers', in Gilhooly, K. J., Keane, M. T. G., Logie, R. H. and Erdos, G. (eds), *Lines of Thinking: Reflections of the Psychology of Thought*, Vol. 1 (pp. 109–24), Chichester: Wiley, 1990.

Olsen, R. A. and O'Neill, M. F., 'The interpretation of probabilistic phrases used to provide financial advice', *Journal of Professional Services Marketing*, **4** (1988), 67–74.

Rapoport, A., Wallsten, T. S. and Cox, J. A., 'Direct and indirect scaling of membership functions of probability phrases', *Mathematical Modeling*, **9** (1987), 397–417.

Saaty, T. L., *An Eigenvalue Allocation Model for Prioritization and Planning*, Energy Management and Policy Center, University of Pennsylvania, 1972.

Saaty, T. L., 'A scaling method for priorities in hierarchical structures', *Journal of Mathematical Psychology*, **15** (1977a), 234–81.

Saaty, T. L., 'Modeling unstructured decision problems: a theory of analytical hierarchies', *Proceedings of the First International Conference on Mathematical Modeling* (pp. 69–77), 1977b.

Saaty, T. L., *The Analytic Hierarchy Process*, New York: McGraw-Hill, 1980.

Saaty, T. L., *Multicriteria Decision Making: The Analytic Hierarchy Process*, Pittsburgh: RWS Publications, 1990.

Shanteau, J., 'Component processes in risky decision making', *Journal of Experimental Psychology*, **103** (1974), 680–91.

Simpson, R. H., 'The specific meanings of certain terms indicating differing degrees of frequency', *Quarterly Journal of Speech*, **30** (1944), 328–30.

Tanur, J. M., 'On the possible dangers of isolation: Comment on "Quantifying probabilistic expressions" ', *Statistical Science*, **5** (1990) 21–2.

Teigen, K. H. and Brun, W., 'Yes, but it is uncertain: direction and communicative intention of verbal probabilistic terms', *Acta Psychologica*, **88** (1995), 233–58.

Tobler, W. R., 'Smooth pycnophylactic interpolation for geographical regions', *Journal of the American Statistical Association*, **74** (1979), 519–36.

Wallsten, T. S. and Budescu, D. V., 'Comment on "Quantifying probabilistic expressions" ', *Statistical Science*, **5** (1990), 23–6.

Wallsten, T. S., Budescu, D. V., Rapoport, A., Zwick, R. and Forsyth, B., 'Measuring the vague meanings of probability terms', *Journal of Experimental Psychology: General*, **115** (1986), 348–65.

Wallsten, T. S., Budescu, D. V. and Zwick, R., 'Comparing the calibration and coherence of numerical and verbal probability judgment', *Management Science*, **39** (1993a), 176–90.

Wallsten, T. S., Budescu, D. V., Zwick, R. and Kemp, S., 'Preferences and reasons for communicating probabilistic information in verbal or numerical terms', *Bulletin of the Psychonomic Society*, **31** (1993b), 135–8.

Weiss, E. N. and Rao, V. R., 'AHP design issues for large-scale systems', *Decision Sciences*, **18** (1987), 43–61.

Wolf, C. Jr, 'Comment on "Quantifying probabilistic expressions" ', *Statistical Science*, **5** (1990), 31–2.

Zahedi, F., 'The analytical hierarchy process — a survey of the method and its applications', *Interfaces*, **16** (1986), 96–108.

Authors' biographies:

Madjid Tavana is Chairman of the Management Department at La Salle University where he has also served as the Director of the Center for Technology and Management. Dr Tavana has been a Visiting Scholar at the Anderson Graduate School of Management at UCLA and a Faculty Fellow in Aeronautics and Space Research at NASA — Kennedy Space Center. His research interests are decision support systems, multi-criteria decision making, and model management.

Dennis T. Kennedy is an Associate Professor of Accounting at La Salle University. Dr Kennedy's research interests are decision analysis, conceptual framework of accounting, and accounting education.

Barbara Mohebbi is an Assistant Professor of Business Mathematics at Fairleigh Dickinson University. Professor Mohebbi's research interests are decision making, total quality management, and health care administration.

Authors' addresses:

Madjid Tavana, Chairman, Management Department, La Salle University, Philadelphia, PA 19141, USA.

Dennis T. Kennedy, Accounting Department, La Salle University, Philadelphia, PA 19141, USA.

Barbara Mohebbi, Edward Williams College, Fairleigh Dickinson University, Hackensack, NJ 07601, USA.

APPENDIX A: PROBABILISTIC PHRASES QUESTIONNAIRE 1

INSTRUCTIONS

It has been claimed in both applied decision analysis and in basic research that people can better use and understand probabilistic opinions expressed by probabilistic phrases than by numeric probabilities. The purpose of this study is to examine the numeric probabilities which individuals associate with verbal expressions.

Would you help us with this study by circling, in each of the following groups of phrases, one verbal expression whose meaning is most representative of the group.

Uncertain	Perhaps
Inconclusive	Not Certain
A Certain Hope	Possible
Possibly	Unpredictable

Improbable	Rare
Not Likely	Small Possibility
Barely Possible	Very Doubtful
Highly Improbable	Not Probable
Small Probability	Rarely
Quite Unlikely	Very Unlikely

Rather	Slightly More Than Half the Time
Better Than Even	Predictable
Somewhat Likely	Slight Odds in Favour

Somewhat Unlikely	Somewhat Doubtful

Probable	Frequently
Pretty Good Chance	Good Hope
Fairly Likely	Probably
Common	Rather Likely
Likely	Often
Good Chance	Predictable

No Chance	Never
Not Possible	Impossible

Quite Likely	Usually
Most Possibly	Great Chance
Very Good Hope	Quite Certain

Small Chance	Improbable
Small Likelihood	Not Very Probable
Seldom	Rather Unlikely
Fairly Unlikely	Quite Doubtful
Uncommon	Unlikely
Doubtful	Usually Not
Not Much Chance	A Small Hope

Always	Certain

Fair Chance	Toss-up
Small Doubt	Slight Odds Against
Fighting Chance	Danger (Risk)
Slightly Less Than Half the Time	

Very Likely	Very Probable
Highly Probable	

APPENDIX B: PROBABILISTIC PHRASES QUESTIONNAIRE 2

INSTRUCTIONS

It has been claimed in both applied decision analysis and in basic research that people can better use and understand probabilistic opinions expressed by probabilistic phrases than by numeric probabilities. The purpose of this study is to examine the numeric probabilities which individuals associate with verbal expressions.

Would you help us with this study by ranking the following verbal expressions from 1 to 11 with 1 assigned to the expression with the highest probability of occurrence and 11 to the expression with the lowest probability of occurrence.

Verbal expression	Probability of occurrence 1 = Highest to 11 = Lowest
Somewhat Doubtful	
Impossible	
Toss-up	
Quite Certain	
Small Possibility	
Somewhat Likely	
Likely	
Certain	
Possible	
Very Likely	
Small Chance	

APPENDIX C: PROBABILISTIC PHRASES QUESTIONNAIRE 3 (PARTS 1 and 2)

INSTRUCTIONS: Please respond by placing an 'X' on the appropriate location to indicate your perception of the relative likelihood of the verbal expressions for each of the following pairwise comparisons.

Placing an 'X' at equal indicates that verbal expressions A and B imply an equal probability of occurrence

Column scale (left to right): Extreme, Very strong, Strong, Moderate, Equal, Moderate, Strong, Very strong, Extreme

Verbal Expression A		Verbal Expression B

Placing an 'X' to the left of equal indicates that verbal expression A implies a higher probability of occurrence than B

Placing an 'X' to the right of equal indicates that verbal expression B implies a higher probability of occurrence than A

PART 1

Verbal Expression A	Verbal Expression B
Possible	Toss-up
Small Possibility	Impossible
Impossible	Small Chance
Possible	Impossible
Small Possibility	Somewhat Doubtful
Somewhat Doubtful	Toss-up
Possible	Somewhat Doubtful
Small Possibility	Small Chance
Somewhat Doubtful	Small Chance
Impossible	Somewhat Doubtful
Small Possibility	Possible
Impossible	Toss-up
Small Possibility	Toss-up
Possible	Small Chance
Small Chance	Toss-up

PART 2

Verbal Expression A	Verbal Expression B
Very Likely	Likely
Somewhat Likely	Toss-up
Certain	Quite Certain
Very Likely	Quite Certain
Quite Certain	Toss-up
Likely	Certain
Very Likely	Somewhat Likely
Likely	Quite Certain
Somewhat Likely	Quite Certain
Certain	Somewhat Likely
Very Likely	Toss-up
Certain	Toss-up
Likely	Somewhat Likely
Very Likely	Certain
Likely	Toss-up

APPENDIX D: A MATHEMATICAL SUMMARY OF AHP

Assume that n probabilistic phrases are being considered by a group of experts. The goal is to provide judgment on the relative weight of these probabilistic phrases. The following is a method of deriving, from relative values associated with pairs of phrases, a set of weights to be associated with individual probabilistic phrases.

Let $c_1, c_2, \ldots, c_n$ be the set of n phrases. The quantified judgments on pairs of phrases c_i, c_j are represented by an $n \times n$ matrix:

$$A = (a_{ij}) \qquad (i, j = 1, 2, 3, \ldots, n)$$

If c_i is judged to be equally probable as c_j, then $a_{ij} = 1$
If c_i is judged to be more probable than c_j, then $a_{ij} > 1$
If c_i is judged to be less probable than c_j, then $a_{ij} < 1$

$$a_{ij} = 1/a_{ji} \qquad a_{ij} \neq 0$$

Thus the matrix is a reciprocal matrix (the entry a_{ij} is the inverse of the entry a_{ji}). a_{ij} reflects the relative probability of c_i compared with phrase c_j. $a_{12} = 1.25$ indicates that c_1 is 1.25 times as probable as c_2.

The vector which represents the priorities of the n phrases can be found by computing the normalized eigenvector corresponding to the maximum eigenvalue of the pairwise comparison matrix A. An eigenvalue of the matrix A is defined as λ which satisfies the following matrix equation:

$$Aw = \lambda w$$

λ is a constant which is called the eigenvalue associated with a given eigenvector w. Saaty's theory implies that the eigenvector which indicates the priorities of n phrases is the one associated with the maximum eigenvalue ($\lambda_{\max}$) of the matrix A. Since the sum of the weights should sum to I, the normalized eigenvector is used.

Saaty suggests a measure of consistency for the pairwise comparisons. When the judgments are perfectly consistent, the maximum eigenvalue λ should equal n, the number of phrases that are compared. In general, the responses will not be perfectly consistent, implying that the maximum eigenvalue is greater than n. The larger this eigenvalue, the greater the amount of inconsistency. The deviation from consistency which is called consistency index (CI) is represented by $(\lambda_{\max} - n)/(n - 1)$. Next, the following random index (RI) table for matrices of order 1–10 using a sample size of 100 is used to generate a consistency ratio (CR), which is the ratio of CI to the average RI for the same order matrix. A CI of 0.10 or less is considered acceptable.

n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.58	0.90	1.12	1.32	1.41	1.45	1.49	1.51

APPENDIX E: EXCERPTS FROM INTERNAL MESSAGES AT FSG

Investors have already taken increased interest-rate volatility into consideration and are expecting interest rates to rise. Therefore, it will *probably* take another 125 basis-points rise in interest rates to create the same effect as did the last 75 basis-point rise.

Low funding costs are of primary importance to debt issuers. As interest rates decline, it is *very probable* that notes will be called early, since it is *not very probable* that a satisfactory rate of return will be received in later years.

The continuation of mortgage market steepening effects on the Treasury Curve is *very doubtful*.

Mortgage refinancing and structural housing turnover are the two most important factors contributing to prepayment speeds. Historically, refinancings have had a much greater effect on speeds than turnover. There is a *good chance* for turnover to become a more important component of speed on many mortgage securities since the current refinancing wave is *very unlikely* to continue.

This analysis suggests that there is a *fair chance* for a further decline in interest rates since the refinancing boom is *unlikely* to continue.

There is a *small probability* that these loans will move into the refinancing range. However, once they move into the refinancing range, the seven-year balloon loans *often* begin to prepay faster. Balloon borrowers tend to be in a higher income bracket, more financially astute, and more sensitive to changes in interest rates.

Bonds bought at a premium rely on interest income to generate yield. It is *predictable* that under certain high prepayments and low LIBOR* scenarios, investors in the floating-rate certificates would not fully recoup their investments.

Usually as interest rates increase and prepayments slow, the swap trust certificates *frequently* pick up yield relative to the 1 month LIBOR.*

In the intermediate term, it is *rare* that the duration of the future contracts will display any changes as interest rates increase. However, these changes *possibly* can occur in the long term.

Home Equity Loan borrowers tend to be in a lower credit-quality range than traditional mortgage borrowers and have fewer sources of funding. Prepayments on Home Equity Loans *always* respond much less to interest-rate changes than do prepayments on traditional mortgage loans.

*LIBOR = London Inter-Bank Offered Rate