Comment

Thomas S. Wallsten and David V. Budescu

The summary of earlier studies and new data offered by Mosteller and Youtz regarding numerical conversions of probability phrases are both fascinating and encouraging. The regularity in mean results is good news for researchers seeking to understand the language of uncertainty, because it indicates that theoretical explanations need not invoke constructs representing population differences or changes in linguistic habits over time. However, the regularity should not be taken to suggest that a major codification of the language of probability is a goal to be pursued, for at least four reasons. First, individual differences in the use and understanding of linguistic probability expressions are large, reliable, and probably very resistant to change. Thus, a codification would give only the appearance, but not the reality of consistent usage. Second, probability phrases have vague meanings to individuals. Any attempt to render them precise will of necessity overlook the important semantic role of this vagueness. Third, context effects on the meanings of probability phrases are substantial and probably cannot be eliminated. Finally, there is often a need to communicate not only a best probability estimate, but also information about the amount and nature of supporting evidence. Probability phrases often fill this need in a way that would be difficult if a simple mapping were established between a set of phrases and a set of probability values. The remainder of this note justifies these four claims, discusses their implications, and offers alternative suggestions to those of Mosteller and Youtz.

CONSISTENT INDIVIDUAL DIFFERENCES

It is surprising that Mosteller and Youtz ignore the variability in their own data, as well as that reported in many of the other studies they cite, when suggesting that terms have fairly constant meanings. Indeed, numerous studies have documented that the intrapersonal variability in understanding probability terms is far less than the interpersonal variability, suggesting that group mean values do not well represent individuals. For example, Budescu and Wallsten's (1985) subjects provided numerical equivalents to probability terms or rank ordered the terms on three occasions each separated by at least three weeks. Intrasubject variance in the rank assigned to a given phrase or implied by the numerical assignments was only a fraction of the intersubject variance. Furthermore, individuals' rank ordering of adjacent terms was very consistent over the replications. Thus, for example, at a group level *probable* and *likely* yielded very nearly the same numerical equivalent, but some individuals consistently ranked *probable* above *likely* while others did the reverse. Certain rankings, of course, were agreed to by virtually everyone (such as *unlikely, likely*). Similarly, both Beyth-Marom (1982) and Johnson (1973) found individuals to be relatively consistent in assigning numerical values to phrases, while simultaneously there was considerable variability over subjects.

Equally as important, it is highly doubtful that people can or will change their usage simply because a codification has been established. Data that support this statement were obtained in Experiment I reported by Wallsten, Fillenbaum and Cox (1986). Subjects were primarily National Weather Service (NWS), media, and research meteorologists. It is well known among this group that the NWS has established guidelines, presumably of the sort called for by Mosteller and Youtz, for the use of specified probability terms in precipitation forecasts. If the probability of precipitation is judged to be 0.10 or 0.20 then the qualifier *slight chance* may be used; 0.30, 0.40, and 0.50 forecasts may include the phrase *chance;* 0.60 and 0.70 forecasts may use *likely.* Other terms are not allowed in presenting precipitation probabilities. The respondents were asked to give numerical probability equivalents in medical scenarios for various phrases, including phrases codified by the NWS. The study was aimed at a particular context effect and more will be said about it below. The point for the present is that the locations, ranges and sensitivity to context of the meteorologists' numerical interpretations were no different than those of other people and not influenced by the NWS guidelines. An additional, but unpublished study using only NWS meteorologists yielded the same results. If this one example can be generalized, then the prospect for people giving up their normal understanding of a phrase for an imposed one is not very good.

Thomas S. Wallsten is Professor of Psychology at the University of North Carolina at Chapel Hill. His mailing address is Department of Psychology CB# 3270, Davie Hall, University of North Carolina, Chapel Hill, North Carolina 27599–3270. David V. Budescu is Associate Professor of Psychology at the University of Haifa, Haifa 31999, Israel.

As a final point on individual differences, people seem to vary widely in the probability phrases they include in their active vocabularies. In all of our studies involving open response formats or the selection of a subset of phrases from a larger list, the variety of choices over subjects has been astonishing (Budescu and Wallsten, 1990; Budescu, Weinberg and Wallsten, 1988; Budescu, Zwick, Wallsten and Erev, 1990; Rapoport, Wallsten, Erev and Cohen, 1990)

PERSONALLY VAGUE MEANINGS

It is noteworthy that not only do people differ consistently in the meanings they attribute to probability phrases, but the individual meanings are not precise. Wallsten, Budescu, Rapoport, Zwick and Forsyth (1986) and Rapoport, Wallsten and Cox (1987) developed techniques for representing an individual's understanding of a probability phrase by means of a membership function over the [0,1] probability interval. These functions bear some resemblance to Mosteller and Youtz' acceptability functions, but they are defined and derived for individuals. Various issues necessary to establish the properties and validity of these functions are considered in detail by Wallsten, Budescu, Rapoport, Zwick and Forsyth (1986) and Rapoport, Wallsten and Cox (1987), to which the reader is referred for details. Wallsten, Budescu and Erev (1988) used such functions to predict choice probabilities in an independent task. Of interest here is the fact that the shapes and locations of the functions for a given phrase differ over individuals, but for an individual are relatively consistent over time. It is reasonable to believe, and will be argued below, that this vagueness allows probability phrases to serve important semantic roles, which people would be generally loathe to, and possibly incapable of, giving up. Further, it is probably this vagueness that allows context to have its strong effect on meaning.

CONTEXT EFFECTS

Although Mosteller and Youtz acknowledge that context affects interpretations of probability terms, they minimize its impact. On the contrary, our reading of the literature is that the impact is severe. For example, Hörmann (1983b) has empirically verified the intuition of Leo Crespi reported by Mosteller and Youtz. Hörmann showed strong effects in the expected directions on the interpretations of expressions of quantity of such factors as the size of the object, its nature, and the relation between the sizes of the target and comparison objects. Similar results on the interpretation of quantity have also been shown by Cohen, Dearnley and Hansel (1958) and by Borges and Sawyers (1974).

More to the immediate point, strong context effects on probability phrases have also been documented. Other studies (e.g., Newstead and Collis, 1987; Wallsten, Fillenbaum and Cox, 1986) have followed Pepper and Prytulak's (1974) original demonstration of the importance of perceived base rate on the interpretation of probability phrases. For example, in Experiment 1 of Wallsten, Fillenbaum and Cox (1986), meteorologists judged a probability phrase in the presence of a high relative frequency medical event to imply considerably greater probability than in the presence of a low relative frequency medical event. Experiment 2 in that paper employed college students in a more complete design in which each respondent evaluated each phrase in a high and a low base rate context equated for semantic content. Base rate effects on interpretation were negligible for low probability phrases but substantial for medium and high probability phrases. Reviews of related research have been prepared by Pepper (1981) and Newstead (1988).

Teigen (1988a) has shown that the interpretation of a phrase can be dramatically altered by the number of equally likely possible events. He demonstrated that probability phrases generally thought to be high are assigned to low probabilities when the low values occur as a consequence of there being more than two equally likely alternatives, but not when the low values occur because the evidence favors one of two binary events.

As a final example of a context effect, it is worth mentioning the important difference between being a source or a target of a probability communication. Data from Fillenbaum, Wallsten, Cohen and Cox (1989) and from Budescu and Wallsten (1990) suggest that receivers of probabilistic phrases interpret them to be vaguer and more central (i.e., closer to 0.50) than intended by the sender. In a similar manner, Brun and Teigen (1988) have demonstrated numerous systematic differences in interpretation of probability phrases between physicians and their patients, as well as among other sources and targets of probability communications.

The upshot of all this is that context effects are real and are great. Their impact on communication may not be severe if all parties view the context identically, but can be substantial otherwise. It is a mistake, we believe, to think that phrases can be assigned fixed values.

THE NECESSITY OF VAGUENESS

It seems to us on the basis of the above evidence that a codification of probability phrases will not work. Moreover, on other grounds we will take the stand, probably heretical to many readers of this journal, that in some cases the vagueness implied by probability phrases serves a useful role. Teigen (1988b) has discussed numerous semantic considerations that underlie the use and understanding of probability phrases. The choice of a phrase for an individual often depends not only on the individual's probability estimate, but also on the amount and nature of the supporting evidence, on the nature of the outcome in question relative to the other possible outcomes, and on the range of the probability interval to which the communicator wants to direct attention.

To the degree that information regarding these aspects of the problem is important to the decision maker, this information should be communicated (although not necessarily by means of probability phrases). Consider a hypothetical example presented by Budescu and Wallsten (1987). An individual, deciding whether to undergo a complicated operation and hearing that the survival chances are 90% may make a very different decision if the 90% estimate is based on the surgeon's personal experience with hundreds of operations in that hospital, rather than on a past history of 10 such operations around the world, none performed by the surgeon in question. Clearly, the 90% estimate is more tenuous in the latter than in the former case, and the patient would want to know that. We are not suggesting that a probability phrase should be used to communicate in the latter case and a numerical value in the former case, only that there is more information to be conveyed to the decision maker than a simple numerical estimate.

People do select probability phrases to indicate the nature of the evidence. Teigen (1988b) has demonstrated, for example, that different probability phrases focus attention on opposite sides of the probability interval. Perhaps individuals select these terms differentially according to the direction in which they feel additional data might push the estimates.

In a series of studies, the results of which surprised us very much, we have demonstrated that at least in certain situations people tend on average to make decisions with equivalent expected values given verbal or numerical descriptions of the uncertainties (Budescu and Wallsten, 1990; Budescu, Weinberg and Wallsten, 1988; Erev and Cohen, 1989; Wallsten, 1990). In more recent, still unpublished research, we are beginning to find that this result is not correct on a trial by trial basis, and further, that the nature of the probability information influences how the probabilities and outcome values are traded off in reaching decisions (Erev, Bornstein and Wallsten, 1989; Erev, Gonzalez and Wallsten, 1989). Moreover, in many decision situations, such as those involving lives or catastrophic risks, decision makers do not attempt to operate on an expected value basis, but rather attempt to be sensitive to the nature of the evidence supporting the probability judgments. One may argue about the normative status of such decision behavior, but one cannot argue about its existence.

We are not urging that probability phrases be used instead of probability numbers in certain situations. A good case can be made, much of it based on evidence summarized in this note, that probability phrases should be avoided when decisions are important. What we are arguing is that probability phrases should not be made to appear more precise than they really are, and further, that there are situations in which the nature of the imprecise judgment is useful information in and of itself. As Budescu and Wallsten (1987) have pointed out, communication about uncertainty should always be as precise as possible, but never more precise than warranted by the data. We concur fully with Mosteller and Youtz' worries that probability phrases may not be efficient and useful for communicating judgments about uncertainty in situations where those judgments are important. However, we do not think that a point, or even an interval, numerical equivalence for each of a set of probability phrases will provide the solution to this problem.

How should the available research be used to aid in the communication of uncertainty? First, it is possible that a subset of phrases can be selected whose meanings are more or less agreed upon. Suggestions to this effect have been made before (Bass, Cascio and O'Connor, 1974; Beyth-Marom, 1982). The selection of phrases must depend upon variability as well as central tendency measures, plus a demonstrated insensitivity to context. We are not optimistic that another such endeavor, even based on all these considerations, will be more successful than the previous ones.

More likely to be successful would be an attempt to uncover the various communication roles that probability phrases serve, and to develop numerical techniques for the same purposes. Individuals could be trained in these techniques and encouraged to use them in important situations. In this manner, people could express information about the state of the evidence and the precision of their opinion, as well as their best probability judgment, without a danger of being misunderstood.

ACKNOWLEDGMENT

Preparation of this manuscript was facilitated by Grant BNS-86-08692 from the National Science Foundation.

ADDITIONAL REFERENCES

BASS, B. M., CASCIO, W. F. and O'CONNOR, E. J. (1974). Magnitude estimation of expressions of frequency and amount. *J. Appl. Psychology* **59** 313–320.

BORGES, M. A. and SAWYERS, B. K. (1974). Common verbal quantifiers: Usage and interpretation. *J. Experimental Psychology* **102** 335–338.

BRUN, W. and TEIGEN, K. H. (1988). Verbal probabilities: Ambiguous, context dependent, or both? *Organizational Behavior and Human Decision Processes* **41** 390–404.

BUDESCU, D. V. and WALLSTEN, T. S. (1987). Subjective estimation of vague and precise uncertainties. In *Judgmental Forecasting* (G. Wright and P. Ayton, eds.) 63–82. Wiley, Chichester.

BUDESCU, D. V. and WALLSTEN, T. S. (1987). Subjective estimation of vague and precise uncertainties. In *Judgmental Forecasting* (G. Wright and P. Ayton, eds.) 63–82. Wiley, Chichester.

BUDESCU, D. V. and WALLSTEN, T. S. (1990). Dyadic decisions with numerical and verbal probabilities. *Organizational Behavior and Human Decision Processes.* To appear.

BUDESCU, D. V., WEINBERG, S. and WALLSTEN, T. S. (1988). Decisions based on numerically and verbally expressed uncertainties. *J. Experimental Psychology: Human Perception and Performance* **14** 281–294.

BUDESCU, D. V., ZWICK, R., WALLSTEN, T. S. and EREV, I. (1990). Integration of linguistic probabilities. *Internat. J. Man-Machine Studies.* To appear.

COHEN, J., DEARNLEY, E. J. and HANSEL, C. E. M. (1958). A quantitative study of meaning. *British J. Educational Psychology* **28** 141–148.

EREV, I., BORNSTEIN, G. and WALLSTEN, T. S. (1989). The information reduction assumption and applications of decision theory. Unpublished manuscript.

EREV, I. and COHEN, B. L. (1989). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes.* To appear.

FILLENBAUM, S., WALLSTEN, T. S., COHEN, B. and COX, J. A. (1989). Some effects of available vocabulary and communication task on the understanding and use of non-numerical probability expressions. Unpublished manuscript.

HÖRMANN, H. (1983b). How many are *some, several,* and *a few*? In *Meaning, Use and Interpretation of Language* (R. Baüerle, C. Schwarze and A. von Stechow, eds.). de Gruyter, Berlin.

NEWSTEAD, S. E. (1988). Quantifiers as fuzzy concepts. In *Fuzzy Sets in Psychology* (T. Zetenyi, ed.) 51–72. North-Holland, Amsterdam.

NEWSTEAD, S. E. and COLLIS, J. M. (1987). Context and the interpretation of quantifiers of frequency. *Ergonomics* **30** 1447–1462.

PEPPER, S. (1981). Problems in the quantification of frequency expressions. In *New Directions for Methodology of Social and Behavioral Science* **9**: *Problems with Language Imprecision* (D. Fiske, ed.). Jossey Bass, San Francisco.

RAPOPORT, A., WALLSTEN, T. S. and COX, J. A. (1987). Direct and indirect scaling of membership functions of probability phrases. *Mathematical Modelling* **9** 397–417.

RAPOPORT, A., WALLSTEN, T. S., EREV, I. and COHEN, B. L. (1990). Revision of opinion with verbally and numerically expressed uncertainties. *Acta Psychologica.* To appear.

TEIGEN, K. H. (1988a). When are low-probability events judged to be "probable"? Effects of outcome-set characteristics on verbal probability estimates. *Acta Psychologica* **68** 157–174.

TEIGEN, K. H. (1988b). The language of uncertainty. *Acta Psychologica* **68** 27–38.

WALLSTEN, T. S. (1990). The costs and benefits of vague information. In *Insights in Decision Making: Theory and Applications. A Tribute to the Late Hillel Einhorn* (R. Hogarth, ed.). Univ. Chicago Press, Chicago.

WALLSTEN, T. S., BUDESCU, D. V. and EREV, I. (1988). Understanding and using linguistic uncertainties. *Acta Psychologica* **68** 39–52.

WALLSTEN, T. S., BUDESCU, D. V., RAPOPORT, A., ZWICK, R. and FORSYTH, B. (1986). Measuring the vague meanings of probability terms. *J. Experimental Psychology: General* **115** 348–365.

WALLSTEN, T. S., FILLENBAUM, S. and COX, J. A. (1986). Base rate effects on the interpretation of probability and frequency expressions. *J. Memory and Language* **25** 571–587.

Comment: Representing and Communicating Uncertainty

Robert L. Winkler

1. INTRODUCTION

We live in a world fraught with uncertainties, and these uncertainties are often communicated via qualitative expressions. Since such expressions are lacking in precision, it is helpful to know what different people might mean when they use specific expressions. Therefore, recent studies of quantitative interpretations of linguistic expressions of uncertainty are valuable to anyone who hears or uses such expressions. The paper by Mosteller and Youtz and its predecessors can help us understand how people represent uncertainty and how the process of representing and communicating uncertainty might be improved.

In these comments I focus on some issues that I view as important in the representation and communication of uncertainty. I discuss sources of variability in interpretations of qualitative expressions of uncertainty in Section 2, with emphasis on differences

Robert L. Winkler is James B. Duke Professor, The Fuqua School of Business and Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina 27706.