Research Report

Tactful or Doubtful?

Expectations of Politeness Explain the Severity Bias in the Interpretation of Probability Phrases

Jean-François Bonnefon[1] and Gaëlle Villejoubert[2]

[1]*Laboratoire Travail et Cognition, Centre National de la Recherche Scientifique, Toulouse, France, and*
[2]*Leeds University Business School, the University of Leeds, Leeds, United Kingdom*

ABSTRACT—***When a statement about the occurrence of a medical condition is qualified by an expression of probability, such as the word* possible*, listeners interpret the probability of the condition as being higher the more severe the condition. This severity bias can have serious consequences for the well-being of patients. We argue that the bias is due to a misconception of the pragmatic function served by the expression of probability. The more severe the condition, the greater the chance that the listener construes the expression as a politeness marker rather than as an uncertainty marker. When this misconception does not occur, neither should the severity bias. An analysis of interpretations of probability expressions using a membership-function approach validates this account. We discuss the consequences of this bias for the communication of risk within and outside the medical domain.***

Natural language is a poor tool when it comes to communicating the likelihoods of states of affairs. Numerical probabilities, of course, are a much better medium for communication of uncertainty, which is why most people prefer to be given numerical rather than verbal estimates of likelihoods (Wallsten, Budescu, Zwick, & Kemp, 1993). Unfortunately, numerical estimates are often unavailable, forcing people to base their decisions on phrases as ambiguous as "X is rather likely" or "Y is highly possible."

Not surprisingly, people are prone to a number of biases in their interpretation of such expressions of probability. Prominent among these is the *severity bias*, which is usually demonstrated with health-related material (Fischer & Jungerman, 1996; Franic & Pathak, 2000; Weber & Hilton, 1990). When a probability expression (*probable*, *possible*, *likely*, etc.) qualifies a statement about a patient developing a medical condition or a side effect of some treatment, listeners interpret the probability as higher the more severe the condition. The severity bias can have serious consequences for the well-being of patients. Overestimating the likelihood of a side effect can encourage choice of the wrong treatment. Misunderstanding the likelihood of developing a condition can make communication between doctor and patient frustrating and counterproductive. This potential for damage is further increased by society-wide policies such as the European recommendation to give verbal estimates of the likelihood of side effects (European Commission, 1998). Some experts have asked for the suspension of such policies until research has yielded further insight into the interpretation of probability expressions (Berry, Raynor, & Knapp, 2003). However, only meager understanding has been achieved so far.

Insight into the mechanics of the severity bias has gone no further than Weber and Hilton's (1990) initial explanation that "more severe events may draw attention to potentially higher probability levels, something that might be labeled a 'worry effect' " (p. 788). We propose that the severity bias derives from what Goffman (1967) and Brown and Levinson (1978/1987) have identified as a fundamental mechanism of human social life, *face-work*. All humans project a sense of positive identity and public self-esteem called "face" and are motivated to support their own and other people's face in social interactions. Many actions, called face-threatening acts, can induce a loss of face (e.g., disagreeing with, criticizing, giving orders to, or embarrassing other people). When such an action is performed, the actor is likely to resort to one of many linguistic strategies that mitigate the face threat. Among these strategies is the use of probability expressions, not to communicate degrees of uncertainty, but rather to reduce the impact of face-threatening acts.

Our explanation of the severity bias involves two assumptions. The first assumption is that in addition to communicating likelihood, which is their standard function, probability expressions can be used as face-management devices when they qualify

Address correspondence to Jean-François Bonnefon, Laboratoire Travail et Cognition, Maison de la Recherche, 5 Allées A. Machado, 31058 Toulouse Cedex 9, France, e-mail: bonnefon@univ-tlse2.fr.

face-threatening acts (e.g., "you might be misinformed," "it is possible you will have to pay for my lunch"). There is ample evidence for this claim, both qualitative (Brown & Levinson, 1978/1987) and quantitative. For example, Youmans (2001) reported a rate of 787 probability terms per 20,000 words (from American English speakers), and in 39% of these cases, the probability expression was used for face-management purposes rather than to communicate likelihood. The second assumption is that probability expressions imply high probabilities when they are used for face management. The word *possibly* in "it is possibly going to snow tomorrow" has the function of communicating likelihood, and accordingly denotes a moderate probability of snowfall. However, in "your bad breath is possibly the reason people shun you," *possibly* has nothing to do with uncertainty, and everything to do with face management. In this context, it denotes a high likelihood. We obtained preliminary evidence for this claim in a previous study (Bonnefon & Villejoubert, 2005), in which we found that the terms *possibly* and *probably* denoted higher likelihoods when they qualified criticisms or impositions than when they qualified non-face-threatening contents.

Probability expressions denote the same probability whatever the condition they qualify, as long as they are perceived to perform their likelihood-communication function and as long as all other things are equal (in particular, that the conditions have the same base rate). But when a physician tells a patient that he or she is going to develop a medical condition, the patient's social face is threatened, and the threat is greatest when the condition is most severe. Therefore, the more severe the patient's condition, the more likely a probability expression will be interpreted as a face-management device, rather than as a likelihood-communication device. Increasingly severe conditions increase the number of speakers and hearers who switch to a face-management interpretation of probability expressions. This shift in interpretation, in turn, increases the average probability attached to the expression—hence the severity bias.

METHOD

Participants

Participants were recruited by third-year psychology students at the University of Toulouse, France, as a course requirement. Each student made a list of several men and women who were older than 18 and not studying psychology, randomly drew one man and one woman from his or her list, and asked them to take part in the study. Of the 810 participants in the final sample (401 men, 409 women; mean age = 31.2, SD = 12.8), 21% had completed graduate school, 47% had an undergraduate education, 19% had graduated from high school only, and the remaining 13% had not graduated from high school. The sample included a large proportion of students (39%), but 61% came from a great variety of non-student professions (including 8% who were unemployed).

Materials and Procedure

Data collection focused on the word *possibly*. The numerical interpretation of this word was assessed by eliciting its *fuzzy membership function* (Zadeh, 1965). This membership function assigns to each value of the probability line [0,1] a number that represents the degree of membership of that value in the concept defined by the phrase. Degree of membership is usually expressed as a real number from 0 to 1, such that memberships of 0 denote probabilities that are absolutely not in the concept, and memberships of 1 denote probabilities that are perfect exemplars of the concept. Other values represent intermediate degrees of membership. Membership functions (originally suggested by Wallsten, Budescu, Rapoport, Zwick, & Forsyth, 1986, and Rapoport, Wallsten, & Cox, 1987) provide subtle and rich representations of the meaning of probability expressions, and the use of such functions has been carefully validated in many studies (for reviews, see Budescu & Wallsten, 1995, and Karelitz & Budescu, 2004).

Membership functions were elicited using the multiple-stimuli method introduced in Budescu, Karelitz, and Wallsten (2003). Participants were asked to imagine that their family doctor had announced they would "possibly" develop each of two medical conditions during the year to come. One condition was deafness, and the other was insomnia (order of presentation was counterbalanced). Insomnia and deafness are of similar prevalence in the French population, from which participants were sampled (i.e., both had incidence rates of about 4% in 2000). After having read, "The doctor tells you, you will possibly suffer from insomnia [deafness] soon," participants were asked: "Does the doctor think the probability that you will suffer from insomnia [deafness] soon is . . . ?" This question was followed by the values 10%, 20%, 30%, and so on, up to 100%. Participants provided a judgment for each of the 10 percentages, using a 10-point scale anchored at *absolutely not* and *absolutely*. The task was then repeated with the second medical condition (deafness or insomnia). In addition, participants were asked to say whether developing deafness or developing insomnia was worse news. Individuals who failed to answer (n = 16) or who judged insomnia to be worse news than deafness (n = 131) were filtered out, yielding the final sample of 810 participants. This choice substantially improved the clarity of the analyses.

Finally, participants were asked whether the doctor was qualifying deafness as "possible" because (a) he was not sure it would happen or (b) he wished to announce the news tactfully. The same question was asked with respect to insomnia, to assess in each case which speech function participants believed the probability term was intended to perform.

RESULTS

"Possible deafness" was judged more probable than "possible insomnia." Membership functions were computed by averaging membership judgments across participants. The top panel of

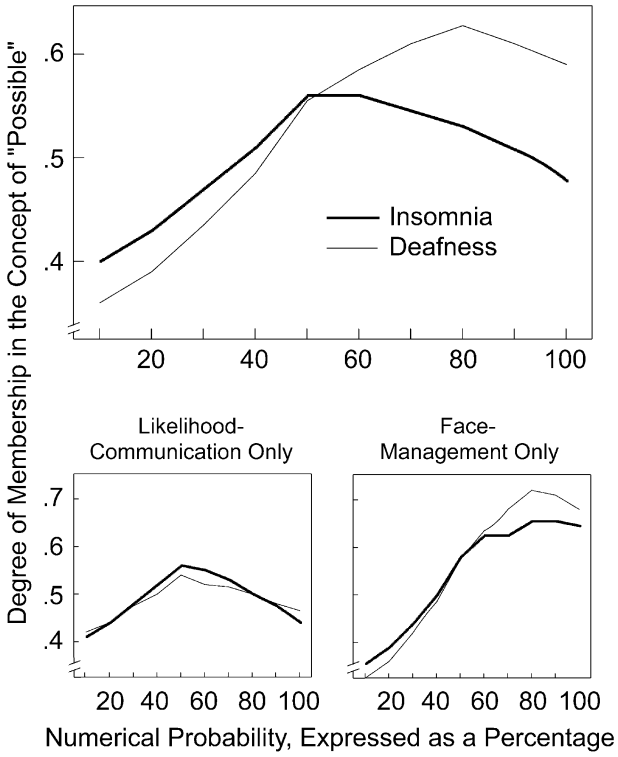

Fig. 1. Membership functions of "possible" deafness and insomnia, for the whole sample (top panel, $n = 810$), for those participants who understood "possible" as serving likelihood-communication purposes (lower-left panel, $n = 669$ for insomnia and $n = 324$ for deafness), and for those participants who understood "possible" as serving face-management purposes (lower-right panel, $n = 141$ for insomnia and $n = 486$ for deafness).

Figure 1 depicts these membership functions overall, without distinguishing which speech function participants assigned to *possible*. These functions show the expected severity bias, as the function for deafness peaks at a higher probability than the function for insomnia. The bias was confirmed by computing the two functions' peaks (see Table 1). The peak for each function was calculated by averaging for each participant the probability values that received the highest membership ratings and then averaging the values obtained across participants.[1] The peak of the function for deafness was higher than the peak for insomnia, $t(815) = 7.90$, $p_{rep} > .999$ (Killeen, 2005), $d = 0.23$ (the 95% confidence interval for this difference was .06–.09).

It appears that one is uncertain about "possible insomnia," but tactful about "possible deafness." As indicated in the last column of Table 1, the proportion of participants who judged that the doctor used the term *possible* for face-management purposes was much greater for deafness (60%) than insomnia (17.4%; $z = 17.1$, $p_{rep} > .999$, Cohen's $h = 0.9$). One noteworthy result is that only 29 participants judged that the doctor was tactful about insomnia but uncertain about deafness. Thus, less than 4% of the participants directly contradicted our hypothesis that more severe conditions increase the likelihood that probability expressions will be perceived as tactful.

[1]It is common for there to be a small discrepancy between the numerical computation of the peak and its graphical representation. Accordingly, the values in Table 1 do not exactly match the peaks of the functions in Figure 1.

TABLE 1

Mean Peaks of the Membership Functions Attached to "Possible" Insomnia and Deafness, as a Function of the Understood Communicative Function of the Probability Term

Medical condition and communicative function	*M* (*SD*)	*n*
Insomnia		
Likelihood communication	.57 (.33)	669
Face management	.71 (.30)	141
Overall	.59 (.33)	810
Deafness		
Likelihood communication	.55 (.34)	324
Face management	.74 (.28)	486
Overall	.67 (.32)	810

The numerical interpretation of *possible* is a function of its being tactful or uncertain. The lower-left panel of Figure 1 depicts the membership functions of *possible* for insomnia and for deafness only for participants who judged that this word communicated likelihood. In contrast, the lower-right panel depicts the same function for participants who judged the term to be tactful, serving face-management motives. The functions within each panel are similar, especially in the case of the lower-left panel (likelihood communication), but the difference between the two panels is dramatic, with the functions in the right panel peaking at higher probabilities than those in the left panel. With respect to insomnia, the peak of the face-management membership function is greater than the peak of the likelihood-communication function (see Table 1 for descriptive statistics), $t(808) = 4.9$, $p_{rep} > .999$, $d = 0.46$ (the 95% confidence interval for this difference was .09–.21). Similarly, with respect to deafness, the peak of the face-management membership function is greater than that of the likelihood-communication function, $t(808) = 8.5$, $p_{rep} > .999$, $d = 0.61$ (with a 95% confidence interval of .14–.23).

Finally, we consider the subsample of participants ($n = 295$) who judged that the doctor was expressing uncertainty about both insomnia and deafness. If politeness expectations do underlie the severity bias, these participants should not have manifested any bias in their numerical interpretations. Indeed, the peaks of the membership functions for insomnia ($M = .57$, $SD = .35$) and deafness ($M = .56$, $SD = .34$) were practically the same for these participants, $t(294) = 0.8$, $p_{rep} = .55$, $d = 0.05$ (the 95% confidence interval for the difference between peaks was −.01–+.03). The two peaks were strongly correlated, $r = .87$.

DISCUSSION

Previous research has established that probability expressions not only communicate degrees of uncertainty, but also serve other pragmatic functions, such as expressing perspective and

drawing attention to the occurrence or the nonoccurrence of the event to which they refer (Sanford & Moxey, 2003; Teigen & Brun, 2003). They may also serve social face-management purposes. We conjectured that the severity bias in interpretations of verbal probability expressions results from the use of such expressions as face-management devices whose function is to safeguard the feelings of people who are receiving face-threatening news.

We tested this conjecture in the medical context of a doctor telling a patient that he or she might develop a mild or severe medical condition. We expected that the face-management interpretation would be more likely when the condition was severe, and that a face-management interpretation would lead to an overestimation of the probability of this condition. Results replicated the severity-bias effect. The same probability word (*possibly*) was judged to communicate higher numerical probabilities when it qualified a more severe condition (deafness) than when it qualified a less severe but equally prevalent condition (insomnia). Furthermore, when the probability term qualified the more severe condition, most participants thought it served a face-management purpose. Also, those who believed the term was used as a face-management device thought that the condition it qualified was substantially more likely to occur than did those who thought the term was communicating a vague likelihood. But participants who believed the probability term was intended to communicate the likelihood of occurrence for both diseases did not exhibit the severity bias.

Thus, people recognize that the more severe a condition is, the more threatening is the news that one has this condition. They also understand that it is polite and tactful to mitigate such face-threatening news by using linguistic moderators, such as probability terms. Finally, they recognize that a probability term used as a face-management device does not refer to the probability of the event it qualifies.

Although our experimental test of this account was limited to the word *possible*, the account should apply to other expressions of probability as well. However, some terms may be less appropriate for expressing politeness than others are. For example, we suspect that *probable* is a less plausible politeness term than *possible*. If so, fewer respondents would interpret information communicated with *probable* as being intended to be tactful. But this would not undermine our main findings. Although a phrase with the term *probable* in it is less likely to be interpreted as serving a face-management purpose than is a phrase with the term *possible*, a statement using *probable* would still be judged to indicate a higher numerical probability when it is interpreted as tactful than when it is interpreted as communicating likelihood information directly (Bonnefon & Villejoubert, 2005). Moreover, when *probable* is interpreted as communicating likelihood, it should receive the same numerical interpretation whatever the severity of the condition to which it refers.

A misunderstanding about which function a probability phrase is intended to serve could lead to a discrepancy between the level of probability a doctor intends to communicate and the level of probability understood by the patient. If the patient interprets a particular phrase as a face-management device but the doctor used it to communicate a vague likelihood, the patient may overestimate the probability the doctor intended. Thus, this research suggests that measures should be taken to ensure that speakers and hearers assign the same communicative function (likelihood communication or face management) to a given probability phrase to improve risk communication.

Acknowledgments—We thank Reid Hastie, Mansur Lalljee, Linda Moxey, and Karl Teigen.

REFERENCES

Berry, D., Raynor, D., & Knapp, P. (2003). Communicating risk of medication side effects: An empirical evaluation of EU recommended terminology. *Psychology, Health and Medicine*, *8*, 251–263.

Bonnefon, J.F., & Villejoubert, G. (2005). Communicating likelihood and managing face: Can we say it is probable when we know it to be certain? In B.G. Bara, L. Barsalou, & M. Buciarelli (Eds.), *Proceedings of the 27th annual conference of the Cognitive Science Society* (pp. 316–321). Mahwah, NJ: Erlbaum.

Brown, P., & Levinson, S.C. (1987). *Politeness: Some universals in language usage*. Cambridge, England: Cambridge University Press. (Original work published 1978)

Budescu, D.V., Karelitz, T.M., & Wallsten, T.S. (2003). Determining the directionality of probability words from their membership functions. *Journal of Behavioral Decision Making*, *16*, 159–180.

Budescu, D.V., & Wallsten, T.S. (1995). Processing linguistic probabilities: General principles and empirical evidence. In J. Busemeyer, D.L. Medin, & R. Hastie (Eds.), *Decision making from a cognitive perspective* (pp. 275–318). San Diego, CA: Academic Press.

European Commission. (1998). *A guideline on the readability of the label and package leaflet of medicinal products for human use*. Brussels, Belgium: Pharmaceuticals Committee, European Commission.

Fischer, K., & Jungerman, H. (1996). Rarely occurring headaches and rarely occurring blindness: Is rarely rarely? *Journal of Behavioral Decision Making*, *9*, 153–172.

Franic, D.M., & Pathak, D.S. (2000). Communicating the frequency of adverse drug reaction to female patients. *Drug Information Journal*, *34*, 251–272.

Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. Garden City, NY: Anchor Books.

Karelitz, T.M., & Budescu, D.V. (2004). You say "probable" and I say "likely": Improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology: Applied*, *10*, 25–41.

Killeen, P.R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, *16*, 345–353.

Rapoport, A., Wallsten, T.S., & Cox, J.A. (1987). Direct and indirect scaling of membership functions of probability phrases. *Mathematical Modeling*, *9*, 397–417.

Sanford, A.J., & Moxey, L.M. (2003). New perspectives on the expression of quantity. *Current Directions in Psychological Science*, *12*, 240–243.

Teigen, K.H., & Brun, W. (2003). Verbal expressions of uncertainty and probability. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgement, and decision-making* (pp. 125–145). Chichester, England: Wiley.

Wallsten, T.S., Budescu, D.V., Rapoport, A., Zwick, R., & Forsyth, B. (1986). Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General*, *115*, 348–365.

Wallsten, T.S., Budescu, D.V., Zwick, R., & Kemp, S.M. (1993). Preference and reasons for communicating probabilistic information in numerical or verbal terms. *Bulletin of the Psychonomic Society*, *31*, 135–138.

Weber, E.U., & Hilton, D.J. (1990). Contextual effects in the interpretation of probability words: Perceived base-rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 781–789.

Youmans, M. (2001). Cross-cultural differences in polite epistemic modal use in American English. *Journal of Multilingual and Multicultural Development*, *22*, 57–73.

Zadeh, L.A. (1965). Fuzzy sets. *Information and Control*, *8*, 338–353.

(Received 9/14/05; Accepted 12/14/05;
Final materials received 1/12/06)