

10. Evidential impact of base rates

Amos Tversky and Daniel Kahneman

In many contexts people are required to assess the probability of some target event (e.g., the diagnosis of a patient or the sales of a textbook) on the basis of (a) the base-rate frequency of the target outcome in some relevant reference population (e.g., the frequency of different diagnoses or the distribution of textbook sales), (b) some specific evidence about the case at hand (e.g., the patient's response to a diagnostic test or the table of contents of the text in question).

Concern with the role of base-rate data in intuitive predictions about individual cases was expressed by Meehl & Rosen (1955), who argued, using Bayes' rule, that predictions of rare outcome (e.g., suicide) on the basis of fallible data is a major source of error in clinical prediction. Meehl & Rosen did not conduct experimental studies but they cited examples from the literature on clinical diagnosis, in which base-rate information was not taken into account.

To obtain an experimental test of the impact of base-rate data, we presented subjects with a description of a graduate student, or a professional, and asked them to predict his field of study or his profession, respectively (Kahneman & Tversky, 1973, 4). These studies showed that posterior probability judgments were determined primarily by the degree to which the description was similar to or representative of the respective professional stereotype (e.g., of librarians or lawyers). The base-rate frequencies of these categories, which were either known to the subjects from their daily experience or stated explicitly in the question, were largely neglected. (We use the term *neglect* to describe situations in which the base rate is either ignored or grossly underweighted.)

Predictions by representativeness or similarity are generally insensitive

This work was supported by the Office of Naval Research under Contract N00014-79-C-0077 to Stanford University.

to base-rate frequencies. However, the phenomenon of base-rate neglect is far more general, since it also occurs in judgments that cannot be readily interpreted in terms of representativeness (Hammerton, 1973). For example, Casscells, Schoenberger, and Grayboys (1978) presented 60 students and staff at Harvard Medical School with the following question:

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person's symptoms or signs? (p. 999)

The most common response, given by almost half of the participants, was 95%. The average answer was 56%, and only 11 participants gave the appropriate response of 2%, assuming the test correctly diagnoses every person who has the disease. Evidently, even highly educated respondents often fail to appreciate the significance of outcome base rate in relatively simple formal problems (see, e.g., Bar-Hillel, 1980a; Lyon & Slovic, 1976). The strictures of Meehl & Rosen (1955) regarding the failure to appreciate base rates are not limited to clinical psychologists; they apply to physicians and other people as well.

The conditions under which base-rate data are used or neglected have been studied extensively by students of judgment and social psychology [see Borgida & Brekke (1981) and Kassin (1979b) for reviews of the literature]. The independent variables investigated in these studies may be divided into two types: procedural and evidential. Procedural variables refer to properties of the design, the task, and the display, while evidential variables refer to the nature of the source and the interpretation of the evidence.

For example, a procedural variable of considerable importance is whether the judge treats each problem as a special case or engages in a task of multiple predictions. Considerable evidence from studies of probability learning and related tasks indicates that people tend to match the distribution of the criterion in making multiple predictions, particularly in the presence of outcome feedback. Because people attempt to generate a pattern of predictions that is representative of the outcome distribution, experiments using repeated judgments with the same base rate produce larger base-rate effects than experiments in which each judgment is treated as a special problem. (See Bar-Hillel & Fischhoff, 1981; Manis et al., 1980).

Another procedural variable of interest is the difference between a within-subjects and a between-subjects design. For example, Fischhoff, Slovic, & Lichtenstein (1979) showed that base-rate data have more impact when the base rates vary in the problems presented to each subject than when different base rates are presented to different subjects. The within-subjects procedure, however, induces a general tendency to assign a higher weight to the varied attribute, even when it is normatively

irrelevant (Fischhoff & Bar-Hillel, 1980). For further discussion of the contrast between comparative (within-subjects) and non-comparative (between-subjects) designs and their implications for the testing of lay statistical intuitions, see Chapter 34.

Although procedural variables have a considerable effect, the present chapter is confined to the discussion of evidential variables that control the interpretation and the impact of the base-rate data. Specifically, we focus on the distinction between two types of base rates, which we label causal and incidental.

Causal and incidental base rates

A base rate is called causal if it suggests the existence of a causal factor that explains why any particular instance is more likely to yield one outcome rather than another. A base rate is called incidental if it does not lead to such an inference.

A compelling demonstration of the contrast between causal and incidental base rates was presented by Ajzen (1977). In one experiment, the respondents assessed the probability that a student, whose academic ability was briefly described, had passed a particular examination. The causal base rate was presented as follows:

Two years ago, a final exam was given in a course at Yale University. About 75% of the students failed (passed) the exam.

This base rate is causal because it implies that the exam was exceptionally difficult (if 75% of the students failed) or relatively easy (if 75% of the students passed). The inferred cause (i.e., the difficulty of the exam) "explains" the base rate and makes every individual student less (or more) likely to pass the exam.

The incidental base rate was presented as follows:

Two years ago, a final exam was given in a course at Yale University. An educational psychologist interested in scholastic achievement interviewed a large number of students who had taken the course. Since he was primarily concerned with reactions to success (failure), he selected mostly students who had passed (failed) the exam. Specifically, about 75% of the students in his sample had passed (failed) the exam.

This base rate is incidental, or non-causal, because the proportion of successful and unsuccessful students in the sample was selected arbitrarily by the investigator. Unlike the causal base rate, it does not permit any inference regarding the difficulty of the exam.

Ajzen's (1977) study showed that the causal base rate was much more potent than the incidental, although variations of both types of base rate produced significant effects. For the causal base rate, the judged probability of success (averaged across descriptions) was higher by .34 when the

base rate of success was high than when it was low. For the incidental base rate, the corresponding difference was only .12. In the terms of the present analysis, the ease or difficulty of an examination is one of the contributing causes that affect the student's performance, and it is therefore integrated with other contributing causes, such as the intelligence and the motivation of the student in question.

The base rate of success was used in the preceding study to define an examination as easy or hard. In a second study, the base rate of preferences was used to define options as more or less attractive (Ajzen, 1977). Subjects were required to assess the probability that students for whom a personality sketch was provided would choose either history or economics as an elective general-interest course. The causal base rate, which served to define the relative attractiveness of the two options, consisted of the proportions of students enrolled in the two courses (.70 and .30). The incidental base rate was introduced as follows:

To obtain student reaction, the history (economics) professor recently interviewed 70 students who had taken his general interest course in history (economics). In order to enable comparisons, he also interviewed 30 students who had taken the course in economics (history).

Note that, unlike the causal base rate, the incidental version provides no information about the popularity of the two courses. The effect of the incidental base rate was not significant in this study, although there was a probability difference of .025 in the expected direction. In contrast, the causal base rate had a strong effect: The mean judged probability of choice was .65 for a popular course (high base rate) and .36 for an unpopular course (low base rate). Evidently, the attractiveness of courses is inferred from the base rate of choices and is integrated with personal characteristics in assessing the probability that a particular student will select one course rather than the other. From a normative standpoint, however, the causal and the incidental base rates in these examples should have roughly comparable effects.

Our next example illustrates a different type of causal base rate; it also permits the calculation of the correct posterior probability under some reasonable assumptions. Consider the following modified version of the cab problem, originally introduced by Kahneman and Tversky (1972a) and later investigated by Bar-Hillel (1980a), and Tversky and Kahneman (1980, 8).

A cab was involved in a hit and run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

(a) 85% of the cabs in the city are Green and 15% are Blue.

(b) a witness identified the cab as Blue. The court tested the reliability of the witness under the same circumstances that existed on the night of the accident and concluded that the witness correctly identified each one of the two colors 80% of the time and failed 20% of the time.

What is the probability that the cab involved in the accident was Blue rather than Green?

To obtain the correct answer, let B and G denote respectively the hypotheses that the cab involved in the accident was Blue or Green, and let W be the witness's report. By Bayes' rule in odds form, with prior odds of 15/85 and a likelihood ratio of 80/20,

$$\begin{aligned} P(B/W)/P(G/W) &= P(W/B)P(B)/P(W/G)P(G) \\ &= (.8)(.15)/(.2)(.85) = 12/17 \end{aligned}$$

and hence

$$P(B/W) = 12/(12 + 17) = .41$$

In spite of the witness's report, therefore, the hit-and-run cab is more likely to be Green than Blue, because the base rate is more extreme than the witness is credible.

A large number of subjects have been presented with slightly different versions of this problem, with very consistent results. The median and modal answer is typically .80, a value which coincides with the credibility of the witness and is apparently unaffected by the relative frequency of Blue and Green cabs.

Base-rate information, however, was utilized in the absence of case data. When item (b) was omitted from the question, almost all subjects gave the base rate (.15) as their answer. Furthermore, the base rate controlled the subjects' expectation about the evidence. A different group of subjects was presented with the above problem except that the sentence "a witness identified the cab as Blue" was replaced by "a witness identified the color of the cab." These respondents were then asked, "What is the probability that the witness identified the cab as Blue?" The median and modal response to this question was .15. Note that the correct answer is $.2 \times .85 + .8 \times .15 = .29$. In the absence of other data, therefore, the base rate was used properly to predict the target outcome and improperly to predict the witness's report.

A different pattern of judgments was observed when the incidental base rate (of cabs) was replaced by a causal base rate (of accidents). This was accomplished by replacing item (a) with

(a') Although the two companies are roughly equal in size, 85% of cab accidents in the city involve Green cabs and 15% involve Blue cabs.

The answers to this problem were highly variable, but the base rate was no longer ignored. The median answer was .60, which lies between the reliability of the witness (.80) and the correct answer (.41). The base rate in (a') is causal because the difference in rates of accidents between companies of equal size readily elicits the inference that the drivers of the Green cabs are more reckless and/or less competent than the drivers of the Blue

cabs. This inference accounts for the differential base rates of accidents and implies that any Green cab is more likely to be involved in an accident than any Blue cab. In contrast, the base rate in (a) is incidental because the difference between the number of Blue and Green cabs in the city does not justify a causal inference that makes any particular Green cab more likely to be involved in an accident than any particular Blue cab.

Note that according to the present analysis the posterior probability that the errant cab is Blue rather than Green is the same under both (a) and (a'). Nevertheless, the correlation between cab color and involvement in accidents is 0 for the incidental base rate and .7 for the causal! This statistical fact reflects the difference between the two base rates and helps explain why the causal base rate is utilized while the incidental base rate is ignored.

Other evidential variables

The causal or incidental nature of base-rate data is not the only evidential variable that affects their impact on intuitive judgments. Even in the absence of a causal interpretation, base-rate data are not superseded by non-specific, impoverished, or incoherent case data. For example, Bar-Hillel (1980a) studied a version of the original cab problem in which the information about the witness (item b) was replaced by a report that the hit-and-run cab was equipped with an intercom and that intercoms are installed in 80% of Green cabs and in 20% of Blue cabs. In this problem, the (incidental) base rate was not discarded, and the median response was .48. Bar-Hillel suggested that the evidence regarding the intercom did not replace the base rate because it is less specific than an identification by a witness. Thus, base-rate data are combined with other evidence either when the former have a causal interpretation or when the latter are no more specific than the base rate (Bar-Hillel, 1980a).

Both specificity and causality may help explain the difference between the results of Kahneman and Tversky (1973, 4), who found an essential neglect of base rate in predicting a student's field of study on the basis of a personality sketch, and the results of McCauley and Stitt (1978), who found a substantial correlation between the judged base rates of traits and the judged probabilities of these traits given a particular nationality, for example, the probability that a person is efficient if he is German. Aside from several procedural differences, the latter study differs from the former in three important aspects. First, subjects were asked to predict relative frequency (e.g., the proportion of Germans who are efficient) rather than the probability for an individual case. Second, the evidence consisted of class membership, for example, German, rather than detailed descriptions of a specific individual. Third, the base-rate frequency of traits may be easier to interpret causally than that of professions. Lay

personality theories suggest reasons why most people are fun loving and only a few are masochistic. These reasons apply to people in general and to Germans in particular, thereby providing a causal interpretation of the base rate of traits.

A situation of special interest concerns specific but non-diagnostic evidence (e.g., a description of a person that is equally similar to an engineer and a lawyer). The experimental findings here are not entirely consistent. Kahneman & Tversky (1973, 4) found base-rate neglect, while Ginosar and Trope (1980) found exclusive reliance on base rate under apparently similar experimental conditions. Most studies, however, obtained intermediate results where the base rate was not discarded but rather diluted by nondiagnostic evidence about the case at hand (see e.g., Manis et al., 1980; Wells & Harvey, 1977).

Internal versus external attributions

A class of base-rate problems of particular interest to social psychologists arises when the evidence and the base rate refer respectively to internal-dispositional and to external-situational factors that affect an outcome. A student's success in an examination, for example, is determined jointly by the difficulty of the exam and by the student's talent. Similarly, one's response to a request to donate money to a particular cause depends on one's generosity and on the nature of the request. External factors, such as the difficulty of an exam or the effectiveness of the request, are naturally expressed by the relevant base rates (e.g., 75% of students failed the exam; most people contributed to the cause). The question regarding the relative impact of situational and dispositional factors in social attribution can thus be reformulated in terms of the weight that is assigned to the corresponding base rates.

Nisbett & Borgida were the first to explore the link between the use of base-rate information in judgment research and the relative weight of situational factors in the study of attribution of behavior. They showed that knowledge of the low frequency of helping behavior in the Darley-Latané (1968) study did not affect subjects' predictions of the behavior of an individual participant in the study, who was observed in a brief filmed interview. The study of Nisbett and Borgida (1975) contributed to the convergence of cognitive and social-psychological approaches to the study of judgment. It also provoked controversy (Borgida, 1978; Wells & Harvey, 1977, 1978) and stimulated a flurry of research on the role of consensus information in the prediction of behavior (Borgida & Brekke, 1981; Kassir, 1979b; Nisbett & Ross, 1980; Ross, 1977).

In contrast to the examples of the exam and the cabs, in which causal and incidental base rates are clearly distinguished, the base rates in many consensus studies are subject to alternative interpretations. To illustrate the point, let us compare the study of Nisbett and Borgida (1975) with the

causal base-rate condition in Ajzen's (1977) experiment, where the subjects evaluated the probability that a particular student passed an exam that 75% of the class had failed. The formal structure of the two problems is precisely the same, but the base rate was largely neglected in the former study and used in the latter. It appears that the surprising base rate was given a situational interpretation in Ajzen's study but was interpreted as an accident of sampling in the Nisbett-Borgida study.

The judgments of Ajzen's subjects indicate that they inferred from the low base rate of success that the exam had been difficult, although they could have used the same evidence to conclude that the students who took the test were inept. In contrast, the subjects of Nisbett and Borgida apparently inferred that the participants in the helping study were mostly unfeeling brutes (Wells & Harvey, 1977). They did not draw the correct conclusion that the situation of the Darley-Latané study is not conducive to helping behavior.

Whether an extreme base rate is attributed to an accident of sampling or to situational factors depends on the context of the problem: It is more plausible that an unusual distribution of test results is due to the difficulty (or ease) of an exam than to the exceptional composition of the class. On the other hand it is harder to revise one's conception about the conditions under which people help a stricken stranger than to assume that the participants in the helping study were exceptionally unhelpful.

The apparent neglect of base-rate data in predictions about individual cases is associated with an inference about unusual characteristics of the members of the group. A causal interpretation of the base rate becomes more likely if this inference is blocked. This hypothesis has been supported by several studies, which restored a base-rate effect by stressing the representativeness of a sample in which surprising behaviors had been observed (Hansen & Donoghue, 1977; Hansen & Lowe, 1976; Wells & Harvey, 1978). The impact of base-rate data was even enhanced in one study by informing the subjects that the sample for which base rates were provided was large and therefore reliable (Kassin, 1979a). The major conclusion of this research is that the use or neglect of consensus information in individual prediction depends critically on the interpretation of that information.