The Psychology of Good Judgment:

Frequency Formats and Simple Algorithms

GERD GIGERENZER, PhD

Mind and environment evolve in tandem—almost a platitude. Much of judgment and decision making research, however, has compared cognition to standard statistical models, rather than to how well it is adapted to its environment. The author argues two points. First, cognitive algorithms are tuned to certain information formats, most likely to those that humans have encountered during their evolutionary history. In particular, Bayesian computations are simpler when the information is in a frequency format than when it is in a probability format. The author investigates whether frequency formats can make physicians reason more often the Bayesian way. Second, cognitive algorithms need to operate under constraints of limited time, knowledge, and computational power, and they need to exploit the structures of their environments. The author describes a fast and frugal algorithm, Take The Best, that violates standard principles of rational inference but can be as accurate as sophisticated "optimal" models for diagnostic inference. *Key words*: Bayes' theorem; bounded rationality; information format; probabilistic reasoning; satisficing; training; medical education. **(Med Decis Making 1996;16:273–280)**

Cognition should not be divorced from its environment, argued Egon Brunswik,[1] comparing the two to a married couple who have to come to terms with one another by mutual adaptation. Judgment and decision making, however, has often been studied as if it were divorced from its environment: by comparing judgment with some statistical rules, and nothing but these rules. On this basis, some researchers have concluded that judgment systematically deviates from statistical rules such as Bayes' rule, emphasizing "cognitive illusions" (e.g., Tversky and Kahneman[2]). To use Brunswik's metaphor, these studies look only at one partner in the couple and try to understand its behavior with respect to a rule, rather than to its partner. The limitations of this approach have recently been documented,[3,4] stirring up a controversy.[5,6] In this paper, I illustrate by two examples how the study of this Brunswikian "couple" can be achieved, and what new light it sheds on diagnostic inference.

I use the term "environment" as shorthand for the *structure* and *format* of the information in an environment. Redundancy, nonlinearity, and stability are examples of structural characteristics of information, and these are indispensable for deciding whether or not some statistical rule is a good norm of sound reasoning. For instance, if one has reason to assume that the environment is not stable but is changing (e.g., a company has been reorganized), and an extreme outcome (e.g., profit) has been observed, then regression to the mean may not be a good norm to predict the next outcome.[7] My first example is the format of information and how it affects Bayesian reasoning; the second is the power of simple, "satisficing" algorithms that exploit the structure of information to make good inferences under constraints of limited time and knowledge.

Received March 8, 1996, from the Center for Adaptive Behavior and Cognition, Max Planck Institute for Psychological Research, Munich, Germany. Accepted for publication March 15, 1996. Presented in part at the 17th annual meeting of the Society for Medical Decision Making, Tempe, Arizona, October 1995, as part of a symposium (together with Kenneth Hammond) on "New Developments in Judgment and Decision Making." Supported in part by DFG grant Ho 1847/1-1.

Address correspondence and reprint requests to Dr. Gigerenzer: Max Planck Institute for Psychological Research, Leopoldstrasse 24, 80802 Munich, Germany;
e-mail: ⟨giger@mpipf-muenchen.mpg.de⟩.

The Format of Information

In a seminal study of how physicians process information about the results of mammography, Eddy[8] gave 100 physicians the following information*:

*For ease of presentation, I use values of 80% and 10% for the sensitivity and false positive rate, respectively, in the study described below. Eddy used slightly different values of 79.2% and 9.6%, respectively.

> The probability that a patient has breast cancer is 1% (the physician's prior probability).
>
> If the patient has breast cancer, the probability that the radiologist will correctly diagnose it is 80% (hit rate or sensitivity).
>
> If the patient has a benign lesion (no breast cancer), the probability that the radiologist will incorrectly diagnose it as cancer is 10% (false-positive rate).
>
> *Question*: What is the probability that a patient with a positive mammography actually has breast cancer?

Eddy reported that 95 of 100 physicians estimated the probability of breast cancer after positive mammography to be about 75%. If one inserts the numbers into Bayes' theorem, however, one gets a value of 7.5%, that is, an estimate one order of magnitude smaller. Casscells and colleagues[9] have reported similar results with physicians, staff, and students at the Harvard Medical School. Is there something systematically wrong with physicians' statistical training, with their intuitions, or both?

Physicians are no exception in having difficulties with probabilities. Numerous undergraduates sitting through tests in psychological laboratories found themselves similarly helpless and were diagnosed as suffering from "cognitive illusions." From these studies, many have concluded that the human mind lacks something important: "People do not appear to follow the calculus of chance or the statistical theory of prediction"[10,p.237]; "It appears that people lack the correct programs for many important judgmental tasks"[11]; or more bluntly, "Tversky and Kahneman argue, correctly I think, that our minds are not built (for whatever reason) to work with the rules of probability."[12,p 469] If these conclusions are correct, then the problem is not so much in training, but in our minds: there seems to be little hope for physicians, and for their patients as well.

MENTAL COMPUTATIONS DEPEND ON INFORMATION FORMATS

These conclusions, however, are premature. Let us be clear why. A discrepancy between human judgment and the outcome of Bayes' rule is observed, from which the conclusion is drawn that there is no cognitive algorithm similar to Bayes' rule in people's minds (but only dubious heuristics such as "representativeness"). However, any claim against the existence of an algorithm, Bayesian or otherwise, is impossible to evaluate unless one specifies the *information format* for which the algorithm is designed to operate. For instance, numbers can be represented in various formats: Arabic, Roman, and binary systems, among others. My pocket calculator has an algorithm for multiplication that is designed for Arabic numbers as the input format. If I enter

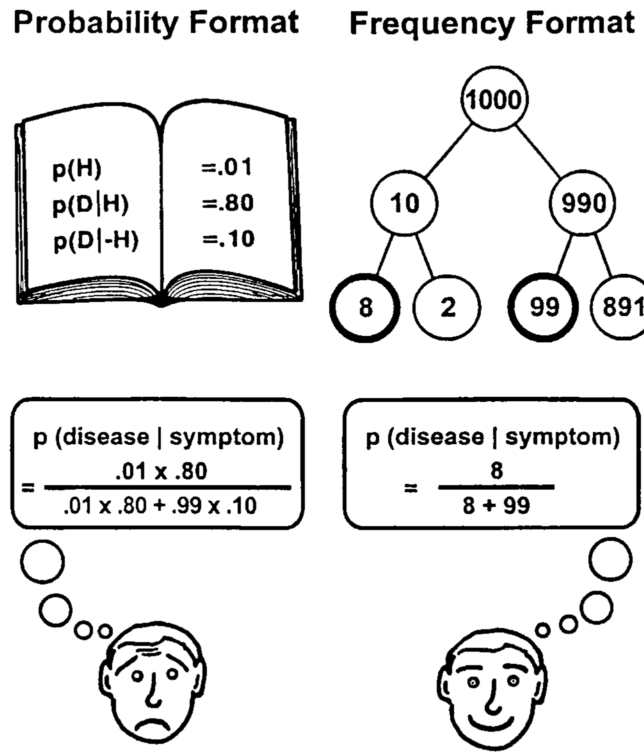

FIGURE 1. Bayesian computations are simpler when information is represented in a frequency format (right) than when it is represented in a probability format (left). p(H) = prior probability of hypothesis. H (breast cancer), p(D|H) = probability of data D (positive test) given H, and p(D|−H) = probability of D given −H (no breast cancer).

binary numbers instead, garbage comes out. The observation that the output of my pocket calculator deviates from the normative rule (here: multiplication), however, does not entail the conclusion that it has no algorithm for multiplication. Similarly, the algorithmic operations acquired by humans are designed for particular formats. Consider for a moment division in Roman numerals.

The format of information is a feature of the decision maker's environment. Let us apply this argument to medical diagnosis, such as Eddy's mammography problem. Assume that through the evolutionary process of adapting to risky environments, some capacity or cognitive algorithm for statistical inference has evolved. For what information format would such an algorithm be designed? Certainly not probabilities and percentages—as in the above mammography problem—because these are relatively new (a few hundred years old) formats for learning and communicating risk.[3,13] So if not probabilities and percentages, for what information format were these cognitive algorithms designed? I argue that they evolved to deal with absolute frequencies, because information was experienced during most of the existence of *Homo sapiens* in terms of discrete cases, for example, three out of 20 cases rather than 15%.

This evolutionary speculation combines with an important result: Bayesian computations are simpler with absolute frequencies than with probabilities or percentages, as illustrated in figure 1.[14,15] On the left side, the information is represented in terms of probabilities, as by Eddy[8] and Casscells et al.,[9] and as in most medical and statistical textbooks. The formula on the left side (Bayes' rule) shows the mental computations needed to estimate the probability of breast cancer after positive mammography. On the right side, in contrast, information is represented in absolute frequencies, as actually learned through experience. Imagine a physician in an illiterate society. She has no books or statistical surveys and must rely solely on her own experience. Her people have been afflicted by a previously unknown disease, and she was lucky to discover a symptom that signals the disease, although this symptom is not a certain predictor. She has seen 1,000 people, 10 of whom had the disease. Of those, eight showed the symptom. Of the 990 not afflicted, 99 also showed the symptom. Now a new patient appears. He has the symptom. What is the probability that he has the disease? The physician does not need a pocket calculator to find the Bayesian estimate. All she needs to do is to relate the number of cases that had both the symptom and the disease (8) to the number of symptom cases (99 + 8). The chances are 8 out of 107, or roughly 8%. Whatever the exact number is, our physician cannot be as easily fooled as her contemporary colleagues into believing that the probability is around 75%.

The formula on the right side (Bayes' rule) is computationally simpler than the one needed for probability information, that is, a smaller number of cognitive operations need to be performed. Henceforth, I use the term "frequency format" to refer to absolute frequencies as defined by the tree in figure 1. Notice that Bayesian computations are not facilitated by all kinds of frequencies (e.g., normalized frequencies such as 10 of 1,000, 800 of 1,000, and 100 of 1,000 would not work), only by absolute frequencies as they would be actually experienced by "natural sampling" (that is, without normalizing[14]). Figure 1 illustrates two related results: only two pieces of information, the symptom-and-disease frequencies and the symptom-and-no-disease frequencies (the two bold circles in figure 1) need to be attended to in frequency formats, and as a consequence, the base-rate frequency (10 out of 1,000) can be ignored.

HOW TO IMPROVE BAYESIAN REASONING

Thus we have an interesting result: It is easier to reason the Bayesian way with frequency formats. Does this stand up to an empirical test? Ulrich Hoffrage and I have investigated whether students unfamiliar with Bayes' theorem can be made to reason the Bayesian way by using frequency formats.[14] In 15 different inferential problems, including the mammography problem, Bayesian reasoning went up from 16% in the probability format to 46% and 50% in two versions of the frequency format. No instruction or feedback was given; the information format by itself improved Bayesian reasoning. Similar results were obtained by Christensen-Szalanski and Beach[16] and Cosmides and Tooby.[17]

Frequency formats can also be used in instruction: as tools in the classroom or in textbooks to communicate information in a more effective way than traditional probability formats. For instance, Peter Sedlmeier and I have designed a computerized tutorial program that instructs people to translate probabilities into frequency formats.[18] Compared with the conventional method of teaching people how to insert probabilities into Bayes' theorem (for control, this method was also taught by a computerized tutorial system), the median performance was doubled when frequency formats were taught. Equally important, performance showed no loss after a five-week interval, whereas conventional teaching showed the typical steep decay curve in students' performances. The frequency-format tutorial took less than two hours.

HOW TO IMPROVE BAYESIAN REASONING IN PHYSICIANS

When I was invited to speak to the Society for Medical Decision Making (the lecture on which this article is based), I thought I should do some homework. I decided to test whether these findings applied to physicians. Can we improve Bayesian reasoning in physicians by communicating information in frequencies instead of probabilities? One might suspect that this method works only with students without experience in diagnostic inference, but not with physicians, who make diagnostic inferences every day. On the other hand, medical textbooks typically present information about sensitivity, specificity, and priors in probability formats (as in figure 1, left side). Physicians may be stuck—like my pocket calculator—fed with numerals for which their minds were not designed.

Ulrich Hoffrage and I conducted a study to look at this question.[19] We used four medical-diagnosis problems: inferring the chances of breast cancer from a mammography result (as in Eddy's work), inferring Bechterev's disease on the basis of HL antigen B 27, inferring phenylketonuria from a Guthrie test result, and inferring colon cancer on the basis of a Hemoccult test result. We consulted experts to determine the best statistical information available for the base rates, sensitivity, and specificity. There

Table 1 • Frequency Format and Probability Format of the Mammography Problem

To facilitate early detection of breast cancer, women are encouraged from a particular age on to participate at regular intervals in routine screening, even if they have no obvious symptoms. Imagine you conduct in a certain region such a breast cancer screening using mammography. For symptom-free women aged 40 to 50 who participate in screening using mammography, the following information is available for this region:

[*Probability format*]

The probability that one of these women has breast cancer is 1%. If a woman has breast cancer, the probability is 80% that she will have a positive mammography test. If a woman does *not* have breast cancer, the probability is 10% that she will still have a positive mammography test. Imagine a woman (aged 40 to 50, no symptoms) who has a positive mammography test in your breast cancer screening. What is the probability that she actually has breast cancer? _____ %

[*Frequency format*]

Ten out of every 1,000 women have breast cancer. Of these 10 women with breast cancer, 8 will have a positive mammography test. Of the remaining 990 women *without* breast cancer, 99 will still have a positive mammography test. Imagine a sample of women (aged 40 to 50, no symptoms) who have positive mammography tests in your breast cancer screening. How many of these women do actually have breast cancer? _____ out of _____

were two versions for each of the four diagnostic problems: a probability format and a frequency format. Table 1 shows the two versions of the mammography problem.

I had never studied physicians before, and it was an experience. For instance, an ear, nose, and throat specialist who was also a university professor was completely beside himself and simply refused to give any numerical estimates: "On such a basis one can't make a diagnosis. Statistical information is one big lie."

We convinced 48 physicians in Munich to participate in the study and provide estimates. Their mean age was 42 years, and the average professional experience was 14 years. Each physician worked on all four problems, two in the probability format and two in the frequency format. Each problem was on one sheet, followed by a separate sheet where the physician was asked to make notes, calculations, or drawings. I describe first the performance of Dr. A., who represents in several respects the average result.

Dr. A. is 59 years old, director of a university clinic, and a dermatologist by training. He spent 30 minutes on the four problems and another 15 minutes discussing the results with the interviewer. Like many physicians, he became visibly nervous when working on the problems, but only when faced with the probability formats. Dr. A. first refused to write notes, later agreeing to do so, but only on his own piece of paper, not on the questionnaire, and he did not let the interviewer see his notes.

Dr. A. first got the mammography problem in the probability format (table 1) and commented, "I never inform my patients about statistical data. I would tell the patient that mammography is not so exact, and I would in any case perform a biopsy." He estimated the probability of breast cancer after a positive mammography result to be 80% + 10% = 90%. That is, he added the sensitivity to the false-positive rate (this is an unusual strategy). Nervously, he remarked: "Oh, what nonsense. I can't do it. You should test my daughter, she studies medicine." Dr. A. was as helpless with the second problem, Bechterev's disease, in a probability format. Here he estimated the posterior probability by multiplying the base rate by the sensitivity (a common strategy of statistically naive students[14]).

After Dr. A. had seen the first problem in a frequency format, his nervousness subsided. "That's so easy," he remarked with relief, and came up with the Bayesian answer, as he did with the other problem in the frequency format. Dr. A.'s reasoning turned Bayesian when information was in frequencies, despite the fact, as he told us, that he did not know Bayes' theorem.

Incidentally, Dr. A. was not the only one who, in despair, referred to his daughter or son. In one case, the daughter was actually nearby and took the test, too. Her father, a 49-year-old private practitioner, had worked for 30 minutes on the four problems and failed on all. "Statistics is alien to everyday concerns and of little use for judging individual persons," he declared. He derived his numerical estimates from one of two strategies: base-rate only, or sensitivity only (again, strategies used by statistically naive students[14]). His 18-year-old daughter solved all four problems by constructing Bayesian trees (as in figure 1). When she learned about her father's strategies, she glanced at him and said: "Daddy, look, the frequency problem is not hard. You couldn't do this either?" For him, even frequency formats didn't help. In contrast, a 38-year-old gynecologist faced with the mammography problem in the frequency format exclaimed: "A first grader could do that. Wow, if someone couldn't solve this. . . !"

GENERAL RESULTS

The 48 physicians worked on the four problems about a half hour on average. When the information was presented in a probability format, the physicians reasoned the Bayesian way in only 10% of the cases. When the information was presented in a frequency format, this number increased to 46%. The physicians spent about 25% more time on the probability problems, which reflects that they found these more difficult to solve. As the case of Dr. A. illustrated, physicians often reacted—cognitively, emotionally, and physiologically—differently to probability and frequency formats. The physicians were more often nervous when information was presented in probabilities, and they were less skeptical of the relevance of statistical information to medical diagnosis when the information was in frequencies. As the various references to daughters and sons indicated, Bayesian responses were age-correlated: The older half of the physicians (more than 40 years old) contributed only 37% of the Bayesian solutions, the younger 63%.

Physicians are often reported to get uneasy or even angry when asked for statistical information,[20] and to believe that their patients do not understand, or do not want to understand, the uncertainties inherent in diagnosis and therapy.[21] I imagine that frequency formats might help improve the communication between patients and physicians[22] and provide a tool for helping the patient to become a more apt decision maker.

Reasoning the Fast and Frugal Way

So far, I have dealt with a specific diagnostic situation with only one piece of data (e.g., a positive mammography result). However, when there are multiple pieces of information (e.g., the results of a series of tests) that are not independent but redundant, Bayes' rule and other "rational" algorithms, such as weighted linear models, quickly become mathematically complex and computationally intractable—at least for a human mind. One way to deal with such situations is to design sophisticated diagnostic systems that combine all available information in some "optimal" fashion. These models can require extensive computational equipment beyond the human mind. The fiction is that of a "Laplacean demon"—a computationally omnipotent and almost omniscient being, with unlimited time, knowledge, and computational power. Theories of mind are populated with variants of this fiction: multiple regression, Bayesian models, neural networks, and others.[23] Every bit of information is taken into account and integrated in some "optimal" fashion. Brunswik and neo-Brunswikians have often

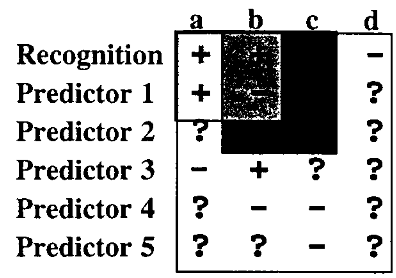

FIGURE 2. Illustration of the Take The Best algorithm. The values of the five ecological predictors are positive (+), negative (–), or unknown (?). Predictors are ordered according to their validities (except for recognition). To infer whether $a > b$, the Take The Best algorithm looks up only the information in the lighter-shaded spaces; to infer whether $b > c$, the search is bounded to the darker-shaded spaces. The other predictor values are not looked up.

subscribed to this view, too, choosing multiple linear regression as a first approximation of how people, including clinical judges, infer properties of their environments.[24–26] But beginning in Brunswik's own writings, one can sense concern about the psychological reality of linear multiple regression.[27]

Must humans try to turn into Laplacean demons, or else be doomed to make bad inferences and decisions? I do not believe that the choice is between "optimal," computationally expensive statistical procedures on one hand and "irrational" heuristics and biases on the other. There is a third way to understand judgment, most prominently represented by Herbert Simon's[28] notion of "satisficing." A satisficing algorithm is not "optimal" in the sense that it searches for all information and integrates it in some optimal way; it is a psychologically plausible strategy that actually can be performed with the limited time, knowledge, and computational strength that real physicians have. Properly designed, a satisficing algorithm can exploit the structure of an environment and perform reasonably well. In what follows, I describe a satisficing algorithm for choice.

TAKE THE BEST

Consider a treatment allocation problem: A choice has to be made between treating one of two patients (or potential patients). The criterion is some benefit to the patient, and there are a number of predictors, such as test results and the patient's age. *Take The Best* is a satisficing algorithm designed for problems of this kind, that is, situations in which inferences have to be made quickly about which of two objects (patients, alternatives) scores higher on some criterion.[29] The general situation is illustrated in figure 2. There are a number of objects (a, b, c, . . .) and a number of predictors that have binary values (the situation can be generalized to continuous predictors).

I illustrate the logic of Take The Best with a demographic problem that we originally used to study its performance: which of two cities has a larger population? Here, *a* and *b* are two German cities, say Bremen and Heidelberg, and examples for cues that indicate higher population are soccer team (whether or not a city has a team in the major soccer league, the "Bundesliga") and state capital (whether or not a city is a state capital). In addition to these ecological cues, there is a knowledge cue, recognition (whether or not the person has heard of the city). Recognition plays a role only when it is correlated with the criterion, as it is with population. Their predictors are ordered according to their (perceived) validities, with Predictor 1 at the top. The ecological cue values can be positive (a city has a soccer team, which indicates larger population), negative (has no soccer team), or unknown (the person has no information). The task is to infer which city, *a* or *b*, has a larger population.

Models that are traditionally considered to lead to "rational" inferences use all pieces of information (predictor values) and integrate these in some way. Take The Best violates both of these tenets. It operates by limited search, that is, it searches only for part of the information, and it does not integrate any information. Its motto is "take the best, ignore the rest." In order to infer which of two cities has the larger population, Take The Best searches through the predictors, one by one, starting from the top, until it finds the first predictor that discriminates, that is, where one city has a positive predictor value and the other has not. Then search is terminated and the inference is made that the city with the positive value has the larger population. For instance, comparing Bremen (*a*) with Heidelberg (*b*), Take The Best looks up first the recognition values, which do not discriminate, because both are positive. The top-ranking ecological cue, the soccer-team cue (Predictor 1), however, does discriminate. Bremen has a soccer team in the major league, but Heidelberg does not. Search in memory (for cue values) is terminated, and the inference is made that Bremen has the larger population. No other predictor value is looked up. Thus, only four out of 12 values in figure 2 (shaded area) are looked up, and none is integrated. Take The Best is noncompensatory: for instance, the remaining positive values of object *b* in figure 2 cannot reverse the decision made solely on the basis of the higher-ranking Predictor 1. Take The Best is also nonlinear: The predictor that determines the choice can vary from inference to inference. For instance, when comparing *b* and *c*, which are both recognized, the algorithm looks up Predictor 1 and finds that Heidelberg has no soccer team, and the value for *c* is not known. Thus, the soccer cue does not discriminate, because only the combination of a positive value and a nonpositive value (negative or unknown) terminates the search. The values on Predictor 2 are now looked up; Heidelberg has a positive value, and the other city has not. This terminates the search, and the inference is made that Heidelberg is larger. This time, six values had to be looked up. Finally, comparing *a* and *d*, the person has never heard of *d*, which terminates the search, and the inference is made that *a* is larger.

Take The Best is a member of a larger family, the PMM ("probabilistic mental models") family of satisficing algorithms, which are all nonlinear, noncompensatory, and work with the principle we call "one good reasoning," that is, they base inferences on only one predictor as opposed to an integration of several.[29,30]

Although Take The Best seems to reflect what people actually do in many situations under constraints of time and knowledge, its simplicity raises the suspicion that it makes highly inaccurate inferences. How could an inference based on only one predictor be even approximately as good as one based on an integration of all information available? In order to test how accurate Take The Best is, Daniel Goldstein and I set up a competition between Take The Best and five integration algorithms, including multiple regression.[29] The task was to infer which of two cities has the larger population, as described above, but for all German cities with more than 100,000 inhabitants (83 cities), with nine ecological cues as predictors. In order to simulate limited knowledge, we created millions of hypothetical subjects, each of whom had a different amount of knowledge, by replacing actual cue values with unknown values. For each of these subjects, the proportion of correct inferences (e.g., whether Heidelberg is really larger than Bonn) in all possible tests ($83 \times 82/2$ pairs of cities) was determined using Take The Best. Similarly, the proportion of correct inferences was determined for the five linear integration algorithms. The stunning result was that Take The Best matched every one of the competing algorithms in accuracy, including multiple regression, and performed better than some. It was also faster (i.e., it searched for less information in memory) than the other algorithms.

This result is an existence proof that fast and frugal algorithms can be as accurate as computationally expensive algorithms that use more knowledge and time. What we do not yet understand very well are the conditions under which Take The Best and other satisficing algorithms perform as well as they do. What we do know is that the success of Take The Best is *not* due to the obvious reason one cue would be as good as many: a high correlation between cues. The pairwise correlations in the test environment were only in the low and moderate range, with an average absolute correlation of 0.19, and a range between −0.25 and 0.54. A challenge is to find out what structures of information (e.g., nonlinearity, redundancy) Take The Best and other satisficing algorithms exploit, and in which environments they will fail.

There is independent evidence that simple, sequential algorithms can perform as well as traditional statistical techniques. Breiman et al.[31] described an algorithm that classifies heart attack patients into "high-risk" and "low-risk" groups. As in Take The Best, the predictors are ordered, the values of the patients on these predictors are dichotomized, and the algorithm proceeds sequentially. The first predictor is whether the minimum systolic blood pressure over the initial 24-hour observation period was higher than 91 mmHg. If not, then no other information is looked up and the patient is classified as "high-risk." If so, then information about a second predictor is searched, namely whether the patient is more than 62.5 years old. If not, then the search is stopped and the patient is classified as "low-risk." If so, then a third and last predictor is looked up, and the classification obtained. Breiman et al. report that this suspiciously simple algorithm is more accurate than the considerably more intricate standard statistical classification methods. Like Take The Best—which is a choice algorithm—this classification algorithm does not look up most of the predictors for classifying heart attack patients (i.e., the 19 variables measured during the first 24 hours at the San Diego Medical Center, where this research was performed) and does not integrate any information.

Summing Up

The two issues—the role of frequency formats and that of fast and frugal algorithms in human inference—sketched in this paper relate to a larger theme: the opposition between the rational and the psychological. What counts as rational is commonly reduced to the domain of logic and probability theory (Kenneth Hammond[32] refers to this as the "coherence" theory of truth) and psychological explanations are called in when things go wrong. This division of labor is the basis of much of research on judgment and decision making. As one economist from Princeton put it, "either reasoning is rational *or* it's psychological."[3] The insight that cognitive algorithms are designed for specific information formats is one step toward connecting the rational with the psychological and ecological. Satisficing algorithms take a more radical step and dispense with the tenets of classic rationality: no integration of predictors, no compensation between predictors, only limited search, and the occasional violation of transitivity.[29] Principles of classic rationality are replaced by fast and frugal psychological principles. The fact that these principles can lead to inferences as accurately as standard "rational" statistical models do forces us to rethink the very nature of sound reasoning.

Satisficing algorithms such as Take The Best provide a new perspective of Brunswik's lens model. In a "one good reasoning" lens,[29] the first discriminating cue that passes through inhibits any other rays of information from passing through and determines judgment. Remember that multiple regression is not the only possible realization of the notion of vicarious functioning (which Brunswik held to be the most important cognitive principle), and that Take The Best is consistent with Brunswik's principle of cue substitution, although it dispenses with his principle of cue integration. A "one good reasoning" lens can explain a puzzling observation. It has been occasionally reported that physicians have claimed to use several criteria to make a judgment but that experimental tests have showed that they used only one criterion. At first glance, this seems to indicate that those physicians made outrageous claims. But it need not be. If a physician's inference works like Take The Best, then these physicians are correct in saying that they use many predictors, but each decision is made by using only one predictor at any time.

If this paper can inspire some researchers in medical decision making to look more deeply than I could at the implications of the research described here—the tuning of Bayesian reasoning to frequency formats, and the power of satisficing algorithms—then it has fulfilled its purpose.

The author thanks Dan Goldstein, Ulrich Hoffrage, and Anita Todd for their helpful comments about earlier versions of the article.

References

1. Brunswik E. Scope and aspects of the cognitive problem. In: Bruner JS, et al. (eds). Contemporary Approaches to Cognition. Cambridge, MA: Harvard University Press, 1964.
2. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974;185:1124–31.
3. Gigerenzer G. Why the distinction between single-event probabilities and frequencies is relevant for psychology (and vice versa). In: Wright G, Ayton P (eds). Subjective Probability. New York: Wiley, 1994, pp 129–61.
4. Gigerenzer G, Murray DJ. Cognition as Intuitive Statistics. Hillsdale, NJ: Erlbaum, 1987.
5. Kahneman D, Tversky A. On the reality of cognitive illusions. A reply to Gigerenzer's critique. Psychol Rev. 1996;103:in press.

6. Gigerenzer G. On narrow norms and vague heuristics: a rebuttal to Kahneman and Tversky. Psychol Rev. 1996;103:in press.
7. Einhorn HJ, Hogarth RM. Behavioral decision theory: processes of judgment and choice. Annu Rev Psychol. 1981;32:53–88.
8. Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A (eds). Judgment under Uncertainty: Heuristics and Biases. Cambridge, U.K.: Cambridge University Press, 1982, pp. 249–67.
9. Casscells W, Schoenberger A, Grayboys T. Interpretation by physicians of clinical laboratory results. N Engl J Med. 1978; 299:999–1000.
10. Kahneman D, Tversky A. On the psychology of prediction. Psychol Rev. 1973;80:237–51.
11. Slovic P, Fischhoff B, Lichtenstein S. Cognitive processes and societal risk taking. In: Carrol JS, Payne JW (eds). Cognition and Social Behavior. Hillsdale, NJ: Erlbaum, 1976.
12. Gould SJ. Bully for Brontosaurus: Further Reflections in Natural History. New York: Penguin Books, 1992.
13. Gigerenzer G, Swijtink Z, Porter T, Daston L, Beatty J, Krüger L. The Empire of Chance: How Probability Changed Science and Everyday Life. Cambridge, U.K.: Cambridge University Press, 1989.
14. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: frequency formats. Psychol Rev. 1995;102:684–704.
15. Kleiter GD. Natural sampling: rationality without base rates. In: Fischer GH, Laming D (eds). Contributions to Mathematical Psychology, Psychometrics, and Methodology. New York: Springer, 1994, pp 375–88.
16. Christensen-Szalanski JJJ, Beach LR. Experience and the base-rate fallacy. Organ Behav Hum Perform. 1982;29:270–8.
17. Cosmides L, Tooby J. Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition. 1996;58:1–73.
18. Sedlmeier P, Gigerenzer G. How to teach Bayesian reasoning in less than two hours. Manuscript submitted for publication, 1996.
19. Hoffrage U, Gigerenzer G. The impact of information representation on Bayesian reasoning. In: Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, in press.
20. Eddy DM. Variations in physician practice: the role of uncertainty. In: Dowie J, Elstein AS (eds). Professional Judgment. A Reader in Clinical Decision Making. Cambridge, U.K.: Cambridge University Press, 1988, pp 45–59.
21. Katz J. Why doctors don't disclose uncertainty. In: Dowie J, Elstein AS (eds). Professional Judgment. A Reader in Clinical Decision Making. Cambridge, U.K.: Cambridge University Press, 1988, pp 544–65.
22. Bursztajn H, Feinbloom RI, Hamm RH, Brodsky A. Medical Choices, Medical Chances. How Patients, Families, and Physicians Can Cope with Uncertainty. New York: Delta, 1981.
23. Gigerenzer G. From tools to theories: a heuristic of discovery in cognitive psychology. Psychol Rev. 1991;98:254–67.
24. Brehmer B. The psychology of linear judgment models. Acta Psychol. 1994;87:137–54.
25. Dowie J, Elstein AS (eds). Professional Judgment. A Reader in Clinical Decision Making. Cambridge, U.K.: Cambridge University Press, 1988.
26. Hammond KR, Hursch CJ, Todd FJ. Analyzing the components of clinical inferences. Psychol Rev. 1964;71:438–56.
27. Elstein AS, Bordage G. Psychology of clinical reasoning. In: Dowie J, Elstein AS (eds). Professional Judgment. A Reader in Clinical Decision Making. Cambridge, U.K.: Cambridge University Press, 1988, pp 109–29.
28. Simon HA. Models of Bounded Rationality. Cambridge, MA: MIT Press, 1982.
29. Gigerenzer G, Goldstein DG. Reasoning the fast and frugal way: models of bounded rationality. Psychol Rev. 1996, in press.
30. Gigerenzer G, Hoffrage U, Kleinbölting H. Probabilistic mental models: a Brunswikian theory of confidence. Psychol Rev. 1991;98:506–28.
31. Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. New York: Chapman & Hall, 1993.
32. Hammond KR. Functionalism and illusionism: can integration be usefully achieved? In: Hogarth RM (ed). Insights in Decision Making. Chicago, IL: University of Chicago Press, 1990.