THE ESSENTIAL BRUNSWIK

Beginnings, Explications, Applications

Edited by
KENNETH R. HAMMOND
THOMAS R. STEWART

Vicarious Functioning Reconsidered: A Fast and Frugal Lens Model

Gerd Gigerenzer and Elke M. Kurz

OXFORD
UNIVERSITY PRESS
2001

24

Vicarious Functioning Reconsidered: A Fast and Frugal Lens Model

Gerd Gigerenzer and Elke M. Kurz

When Egon Brunswik left Vienna for Berkeley in 1937, he began to abandon his favorite tool—a measurement tool known as the *Brunswik ratio*. The ratio measured the degree of perceptual constancy. In its place, Brunswik (1940b) adopted new tools from the Anglo-American statistical tradition—correlation and regression. Then, he measured the degree to which perception attains the distal stimulus by a correlation coefficient called *functional validity*.

New tools often inspire new theories, a source of new ideas known as the *tools-to-theories heuristic* (Gigerenzer, 1991). Brunswik used this heuristic, and so have many others. After he had switched to correlations as his new tool, his concept of the mind changed, too. He began to regard the mind as an "intuitive statistician," and he suggested that the intuitive statistician would use the same new tools: correlation and regression. In particular, *vicarious functioning*, which Brunswik considered the most fundamental principle of a science of perception and behavior, began to be modeled by multiple regression (as first fleshed out by Hammond, Hursch, & Todd, 1964; see also Tucker, 1964; Stewart, 1976).

Vicarious functioning carried a rich meaning for Brunswik. He agreed with W. S. Hunter that the "flexibility and exchangeability of pathways relative to an end," that is, vicarious functioning, was the defining and unifying criterion of psychology (Brunswik, 1952). Brunswik's classical examples were the substitution mechanisms in psychoanalytic theory, the habit family hierarchy in Hull's behaviorism, and hierarchies of perceptual cues. The psychoanalytic work of his wife, Else Frenkel-Brunswik (e.g., 1942), stressed the fact that one cause can manifest itself in various symptoms—rationalization, hysteric conversation, regression, cathexis, and narcissism, among others. If one symptom is blocked or not available, it can be substituted by another. Similarly, in Hull's habit family hierarchy, if a habit is not successful in a situation, it will be replaced with the next one in the hierarchy. In Brunswik's perceptual research, vicarious functioning had a very specific meaning: It signified the divergent as well as the convergent part of his lens model, the first being ecological validity and the second utilization. Both of these complementary aspects of vicarious functioning were modeled by correlation statistics.

In this chapter, we propose a radically different way to model vicarious functioning: the framework of *fast and frugal heuristics* (Gigerenzer, Todd, & the ABC Research Group, 1999). Simple heuristics are psychologically plausible alternatives to multiple regression, and we argue that they are consistent with Brunswik's own ideas. The adaptive value of vicarious functioning is not only in making accurate judgments, but also in being able to make judgments quickly and with limited knowledge. We illustrate a fast and frugal lens model (there are several, depending on the task and the heuristic) and report a counterintuitive result: In making inferences about real-world criteria (the "distal" stimuli), the fast and frugal lens not only was as accurate but also was even more accurate than the computationally complex multiple regression model.

Vicarious Functioning Reconsidered

The idea of vicarious functioning is an extension of Brunswik's earlier notion of cue learning, which in turn is based on Helmholtz's controversial concept of unconscious inferences and Bühler's duplicity principle (details in Doherty & Kurz, 1996; Gigerenzer & Murray, 1987, pp. 61–81). Vicarious functioning describes adaptive cognitive processes that can handle two constraints: the presence of *uncertainty* and the need for *substitution*. A cue (e.g., the retinal image of an object) is only an *uncertain* indicator of a distal stimulus (e.g., the distance to the object), and a cue may not always be present; thus, an adaptive system has to rely on multiple cues that can be *substituted* for each other.

Is multiple regression an appropriate model of vicarious functioning in all situations? What cognitive processes does it imply, and which does it neglect? We begin by pointing out two cognitive processes implied by the multiple regression model and the evidence that one of these two seems to be dispensable in many situations. Then we will draw attention to two processes inherent in vicarious functioning that multiple regression does not model.

Weighting and Summing?

Two fundamental processes in multiple regression are the weighting of cues and the summing of the cue values (Kurz & Martignon, 1998). Weighting and summing have been used to define rational judgment at least since the Enlightenment—the concepts of expected value and utility, Benjamin Franklin's moral algebra, and *Homo economicus* all rely on these two principles. Why should vicarious functioning not work this way, too?

The first blow was delivered to weighting. In the 1970s and 1980s, Robyn Dawes (e.g., 1979) and his colleagues studied predictive accuracy, that is, situations in which the regression weights were computed from one sample and used to make predictions for a new sample. They showed that simple unit weights, such as +1 and −1, typically led to the same predictive accuracy as the "optimal" weights in multiple regression. Weighting does not seem to matter, as long one gets the sign right. Of course, multiple regression would be more accurate than Dawes's unit weight rule in *fitting* given data (as opposed to *predicting* new data), as models generally do when they have more free parameters. But the purpose of vicarious functioning is to predict what is not yet known rather than to fit what is already known. Thus, the question is: If summing without weighting is as accurate as multiple regression, and much simpler to perform, why should mechanisms of vicarious functioning have evolved that try to estimate regression weights?

The second blow was delivered to summing. During the 1990s, the counterintuitive evidence accumulated that fast and frugal heuristics that do not sum cue values but rely only on the first cue that differentiates between two alternatives can be more accurate than multiple regression (Gigerenzer, Czerlinski, & Martignon, 1999; Gigerenzer & Goldstein, 1996). For instance, the *Take The Best* heuristic (see below) uses a simple form of weighting (namely, ordering cues), but it does not sum the cues. Thus, the question is: If weighting without summing can be as accurate as multiple regression, why should mechanisms of vicarious functioning have evolved that try to sum cue values?

It seems that either weighting or summing is dispensable, but not both. But this is not yet the whole story: Weighting and summing model only a part of vicarious functioning. Two cognitive processes in vicarious functioning are not captured by weighted or unweighted linear models.

Searching and Stopping!

Multiple regression models one of three processes involved in vicarious functioning, the *decision rule* ("judgment policy"), by assuming the use of weighting and summing. It does not model two processes that precede a decision, *rules for search*, which give direction to the search for cues, and *rules for stopping*, which stop this search at some point. Modeling search and stopping is paramount for situations involving limited time, limited knowledge, and other constraints. For instance, when deciding about whom to hire, whom to marry, or which stock to buy, one needs to search for cues—in internal memory or in the external world—and this search cannot go on endlessly. Limited search and stopping rules are the essence of bounded rationality (Simon, 1955), as opposed to the fiction of unbounded rationality. The study of search and stopping rules, however, is bypassed in many experimental designs in which all cues are laid out conveniently in front of a participant, who is not supposed to search for further cues. Convenient packaging, however, does not capture the spirit of representative design, nor that of vicarious functioning. The focus on multiple regression in Brunswikian research has thrown search and stopping out of focus.

In the following, we propose an alternative conception of vicarious functioning that uses weighting but not summing and that models search and stopping. It is a step toward a class of psychologically plausible models of vicarious functioning in human judgment.

A Fast and Frugal Lens Model

The term *fast and frugal* signifies cognitive processes that allow one to make judgments that are reached under limited time and with limited knowledge and that do not try to optimize. *Optimizing* may involve computing the optimal linear weights or the Bayesian conditional probabilities, and *optimizing* has been the classical definition of rationality. Brunswik, however, did not think that the cognitive system is rational, only ratiomorphic or quasi-rational. But he was not clear about the mechanism of these quasi-rational processes. Heuristics that are fast (that is, involve little computation) and frugal (that is, search for only few cues) can define the quasi rationality of Brunswik's lens model.

How would a fast and frugal lens function? It embodies heuristics, principles for search, stopping, and decision. We explain its functioning for two-alternative choice tasks, such as to infer which of two U.S. cities has a higher homelessness rate, or which of two soccer teams will win a game. The specific heuristic we use is the Take The Best heuristic, which is derived from probabilistic mental models theory (Gigerenzer, Hoffrage, & Kleinbölting, 1991). This heuristic is just one illustration; there are other heuristics of similar design and for other tasks, such as for estimation and classification (see Gigerenzer, Todd & the ABC Research Group, 1999). For simplicity, we assume that all cue values are binary (positive or negative, with positive indicating higher criterion values) and ignore the recognition heuristic, the initial step of Take The Best (see Gigerenzer & Goldstein, 1996):

Step 1. Search rule: Choose the cue with the highest validity that has not yet been tried for this task. Look up the cue values of the two objects.

Step 2. Stopping rule: If one object has a positive cue value and the other does not (i.e., either negative or unknown value) then stop search and go on to Step 3. Otherwise go back to Step 1 and search for another cue. If no further cue is found, then guess.

Step 3. Decision rule: Predict that the object with the positive cue value has the higher value on the criterion. (See Figure 24.1.)

This fast and frugal lens uses *one-reason decision making*; that is, the decision is based on only one cue. Take The Best orders cues according to their validities v_i:

$$v_i = \frac{R_i}{R_i + W_i},$$

where R_i is the number of right (correct) inferences, and W_i is the number of wrong inferences based on Cue i alone (among all cases where one object has a positive value and the other does not). Ordering cues according to v_i is fast, but not "optimal." For instance, this order does not try to account for conditional validities of cues, that is, dependencies between cues.

How Accurate Is a Fast and Frugal Lens?

How does the fast and frugal lens compare to the multiple regression lens? We tested four models of vicarious functioning, including Take The Best and multiple regression, in twenty real-world environments. The two other models were Dawes's rule, a linear model that uses unit weights (+1 or −1), as mentioned above, and the Minimalist heuristic, which is like Take The Best except that it is even simpler because it looks up cues in a random order (thus, the only difference is in Step 1). The criteria to be predicted in the twenty environments included economic variables such as selling prices of houses and professors' salaries; psychological variables such as predicting the perceived attractiveness of famous men and women; demographic variables such as mortality rates in U.S. cities and population sizes of German cities; environmental variables such as amount of rainfall, ozone, and oxidants; health variables such as obesity at age eighteen, and sociological variables such as dropout rates in Chicago public high schools. The task was always to predict which of two objects scored higher on a criterion. The data sets ranged from seventeen objects to 395 objects, and from three cues (the minimum to distinguish among the strategies) to nineteen cues (for details see Czerlinski, Gigerenzer, & Goldstein, 1999).

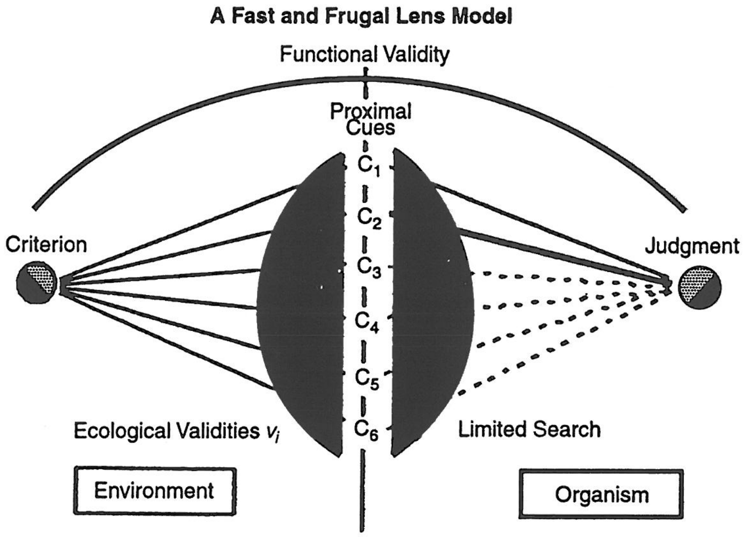

FIGURE 24.1 Illustration of a fast and frugal lens model. The task is to infer which of two objects has a higher value on a criterion. Cues are binary and looked up in the order of their estimated validity. The first cue, C_1, does not discriminate (light line), but the second cue does (dark line). Search is stopped, and the values of C_2 determine the inference. Information concerning other cues is not searched (broken lines).

For each of the twenty tasks, each of the four strategies estimated its parameters (the sign, the order, or the beta weights of the cues) using one half of the objects, and then used these parameters to make predictions about the other half of the objects. This procedure is known as *cross-validation*.

What price does one-reason decision making have to pay for being fast and frugal? Table 24.1 shows how frugal the two heuristics actually were: on average, they searched for fewer than a third of the cues (similar to the illustration in Figure 24.1), whereas the two linear strategies always looked up all cue values (which averaged 7.7 across the twenty environments). How much more accurate were the two linear models than the heuristics? Table 24.1 shows a counterintuitive result: The fast and frugal lens model (using Take The Best) achieved the greatest predictive accuracy, with an average of 71 percent, compared to multiple regression and Dawes's rule with 68 percent and 69 percent, respectively.

This result seems paradoxical because multiple regression processed all the information that Take The Best had, and more. There are two factors that explain this result: the *robustness* and the *ecological rationality* of the fast and frugal lens. The fast and frugal lens is relatively robust, whereas multiple regression overfits. This can be seen from the "fitting" column in Table 24.1. In this condition, Czerlinski, Gigerenzer, and Goldstein (1999) gave all four strategies the complete information about all objects (i.e., no cross-validation), and they fit the data as well as they could. In this case, multiple regression achieved the highest accuracy. The difference between the fitting and the predictive accuracy columns reveals that multiple regression overfitted more than any of the other three strategies.

Dawes's rule lives up to its reputation as a robust strategy, and, consistent with earlier demonstrations, its predictive accuracy matched that of multiple regression. The most frugal strategy, the Minimalist, had to pay some price for simplicity, but not a high one: Its performance was not too far behind that of the two linear strategies in predictive accuracy.

Ecological Rationality

We mentioned overfitting as one reason that the fast and frugal lens performed better than the

TABLE 24.1 Performance of a Fast and Frugal Lens Using a Heuristic (Minimalist or Take the Best) Compared to Two Linear Strategies (Dawes's Rule, Multiple Regression) across Twenty Data Sets

		Accuracy (% correct)	
Strategy	*Frugality*	*Fitting*	*Predictive Accuracy*
Minimalist	2.2	69	65
Take The Best	2.4	75	71
Dawes's rule	7.7	73	69
Multiple regression	7.7	77	68

Note: The average number of cues was 7.7. Performance was measured in terms of frugality (average number of cues looked up) and accuracy (percentage correct). Accuracy was measured both for fitting given data (test set = training set), and for predicting new data, that is, predictive accuracy (test set ≠ training set). The average number of cues looked up was about the same for fitting and generalization (see Czerlinski, Gigerenzer, and Goldstein, 1999).

multiple regression lens. But there is another reason that explains why and when fast and frugal heuristics perform well, even in a purely fitting task (recall that Take The Best was very close to multiple regression even in fitting; see Table 24.1). This second reason is a match between the structure of the (known) information in the environment and that of the heuristic. Brunswik had seen the importance of analyzing the structure of environments in order to understand the mechanisms of the mind, but there is room for improvement in his first attempt to capture the structure in terms of correlation coefficients.

What we call *ecological rationality* is an elaboration of the Brunswikian program of studying the texture of environments. Heuristics are not rational in the classical sense of coherence—the Minimalist, for instance, can produce intransitive judgments. They derive their rationality through a match with the structure of the environment, not with the laws of logic or probability. Martignon and Hoffrage (1999) introduced two characteristics of environments that explain when and why a fast and frugal lens that operates with Take The Best is accurate: *noncompensatory* and *scarce* information.

Noncompensatory Information

The fast and frugal lens is noncompensatory: The decision based on the first cue that discriminates (in the example: C_2) cannot be reversed by the other cues (C_3, C_4, . . .), nor by a combination of them. A noncompensatory set of cues is a set in which each weight is larger than the sum of all other weights to come, such as 1/2, 1/4, 1/8, ... To the extent that cues are noncompensatory, Take The Best will be as accurate as the best linear model. The following theorem states an important property of noncompensatory models and is easily proved (Hoffrage & Martignon, in press).

> *Theorem:* Take The Best is equivalent—in accuracy, not in process—to a weighted linear model whose weights form a noncompensatory set.

If multiple regression happens to have a noncompensatory set of weights (in which the order of this set corresponds to the order of cue validities), then its accuracy is equivalent to that of Take The Best. For instance, among the twenty environments, Martignon and Hoffrage (1999) found four in which this was the case. The important difference between the fast and frugal heuristics and multiple regression or optimization methods is that a fast and frugal lens does not try to compute optimal weights. These heuristics just "bet" that the environment has a structure they can exploit.

Scarce Information

In order to illustrate the concept of scarce information, let us recall an important fact from information theory: A class of N objects contains $logN$ bits of information. This means that if we were to encode each object in the class by means of binary cue profiles of the same length, this length should be at least $logN$ if each object is to have a unique profile. For instance, eight objects can be perfectly predicted by three ($\log 8 = 3$) binary cues. If there were only two cues, perfect predictability simply could not be achieved.

> *Definition:* A set of M cues provides *scarce* information for a reference class of N objects if $M < logN$.

Based on this definition, the following theorem relates the performance of Take The Best to that of Dawes's rule.

> *Theorem:* In the case of scarce information and small numbers of objects (up to 2^{10}), Take The Best is on average more accurate than Dawes's rule.

The proof is in Hoffrage and Martignon (in press). The phrase "on average" means across all possible environments, that is, all combinations of binary entries for NM matrices. The intuition underlying the theorem is the following: In scarce environments, Dawes's rule can take little advantage of its strongest property, namely, compensation. If, in a scarce environment, cues are redundant—that is, if a subset of these cues does not add new information—things will be even worse for Dawes's rule. Take The Best suffers less from redundancy because decisions are made at a very early stage.

The Adaptive Toolbox

We illustrated the mechanism of a fast and frugal lens for a two-alternative choice task. Other types of tasks, such as estimation (Hertwig, Hoffrage, & Martignon, 1999) and classification (Berretty, Todd, & Martignon, 1999), can be performed by heuristics based on similar building blocks that define search, stopping, and decision. This collection of heuristics and their building blocks is what we call the *adaptive toolbox*—specialized mechanisms of cognition and learning that have evolved in the human mind (Gigerenzer, Todd, & the ABC Research Group, 1999). The adaptive toolbox refers to vicarious functioning on the level of heuristics, rather than to cues. An adaptive mind should be able to substitute heuristics just as it does cues.

The specific fast and frugal lens we proposed here embodies limited search and stopping, which the multiple regression model does not incorporate. The fast and frugal lens relies on a simple form of weighting (ordering by validities v_I) but does not use summing or other forms of integrating cue values. Its strength is in its robustness, ecological rationality, and psychological plausibility (Rieskamp & Hoffrage, 1999).

The fast and frugal lens combines Brunswik's ideas of vicarious functioning with the notion of bounded rationality (Simon, 1955). The emphasis on cue substitution as opposed to cue integration is consistent with some of Brunswik's favorite examples: the alternative manifestation and substitution of symptoms in Frenkel-Brunswik's (1942) psychoanalytic work and the substitution of behavior in Hull's hierarchical habit family. And cue substitution is certainly also consistent with an emphasis on the mere rank order of cues (Brunswik, 1947): "In a well-adjusted organism or species, however, the rank order of utilization in what may be called the 'or-assembly' of cues, or the 'cue family hierarchy,' should be the same as the order of their ecological validity"(p. 48). Multiple regression is not the last word on vicarious functioning. We propose taking Brunswik's notion of the quasi-rational nature of vicarious functioning seriously and model it with adaptive heuristics.