UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Inference from Ignorance: The Recognition Heuristic

Permalink

https://escholarship.org/uc/item/5r18d0pq

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 20(0)

Author

Goldstein, Daniel G.

Publication Date

1998

Peer reviewed

Inference from Ignorance: The Recognition Heuristic

Daniel G. Goldstein (dang1@leland.stanford.edu)

Stanford University, Department of Engineering-Economic Systems & Operations Research,
Terman Engineering Center, Stanford, CA 94305-4027 USA

Abstract

While a hindrance to statistical and computational models of inference, missing knowledge can be exploited by organisms in their natural environments. The recognition heuristic utilizes missing knowledge to make accurate inferences about the real world. A consequence of applying this heuristic is a counterintuitive less-is-more effect where less knowledge is better than more for inferential accuracy. Theoretical arguments and experimental evidence supporting the less-is-more effect are given.

The Missing Value Problem?

In statistics, artificial intelligence, and many computational models of mind, missing values are perceived as a nuisance which must be replaced or eliminated before proceeding with the business of inference. In contrast, in interactions between minds and natural environments, patterns of missing knowledge carry information which can serve as the basis for intelligent inference. Consider the eating habits of wild rats which come to recognize foods from their own diet, or from smelling food traces on the breath of conspecifics (Barnett, 1963). These animals have been shown to exhibit a strong *neophobia*, that is, a reluctance to eat foods they do not recognize. This preference can be so strong that rats may prefer recognized foods over novel ones even if they recognize them from the breath of neighbors who appear to be suffering from food poisoning at the time (Galef, 1987; Galef, McQuoid, & Whiskin, 1990). While this behavior seems like a bias, recognition-based food choice helps rats avoid poisoning: any food a living rat or its living neighbors has chosen was clearly not deadly. Missing knowledge, specifically, a lack of recognition, informs food choice in rats. In this paper, I shift focus from animal behavior to human inference and examine the prevalence and consequences of reasoning by recognition in our own species.

We live surrounded by proper names, some which we recognize, and some of which we do not. The names we recognize are not a haphazard collection, they tend to stand out on various criteria which people find interesting. We may notice that the cities whose names we recognize are often larger than those we do not. The corporations we recognize tend to generate more revenue, and possess a greater market share, than those we do not. The scholars we recognize tend to draw larger audiences than those we do not. How do the names we recognize come to reflect such a variety of variables in the world? Names which are outstanding on criteria that people find interesting are more likely than others to be spoken and written about, and are subsequently more likely than others to become recognized. The probabilistic inference can be made that the recognized names are outstanding and unrecognized ones less so.
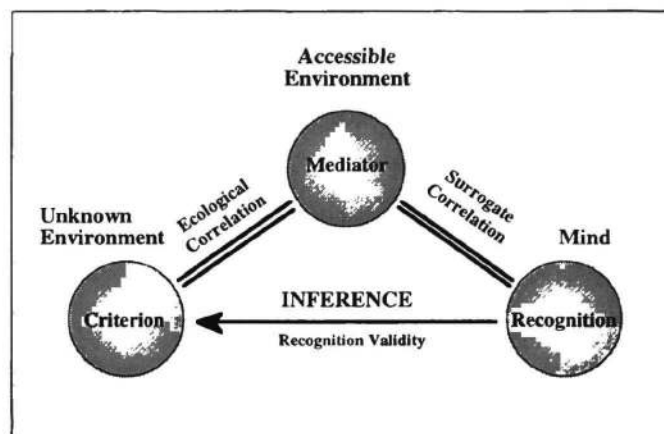

Figure 1: To make inferences about unknown features of the environment, we may exploit the relationship between the presence of items in recognition memory and a criterion. This relationship comes about through mediators in our accessible environment which reflect the criterion.

A probability cue framework, inspired by Brunswik (1955), describes how inferences may be drawn from recognition memory. Figure 1 illustrates how unknown criteria which can not be directly perceived, such as the endowment of a university, or the deadliness of a disease, come to be reflected in the proper names we recognize. For example, let the population of a city be the unknown criterion we wish to infer. Population cannot be directly observed, however, a city's population correlates rather well how often it is mentioned in the newspaper. The newspaper is a *mediator* which casts unobservable aspects of the environment into our immediate surroundings. The correlation between the population of cities and how often they are mentioned in the newspaper is designated the *ecological correlation*. How often cities appear in the newspaper correlates with the number of people who recognize them, denoted by the *surrogate correlation*. Through the two correlations, the mind can infer the unknown criterion. For instance, people may infer that the cities they recognize are larger than the cities they do not recognize. The correlation between the number of people who recognize a city and its population is the *recognition*

correlation. The degree to which inferences based on recognition are accurate is the *recognition validity*. Many approaches to ecological psychology stress the importance of the relationship between the mind, probability cues, and the environment. In the following section, I demonstrate now how these relationships can be quantified in a real-world domain.

Measuring the Relationship Between Memory, Mediator, and Environment

This paper deals with a model domain for inference: the set of cities in Germany with more than 100,000 inhabitants (83 cities in 1994). To measure the strengths of the various correlations in this domain I compared the populations of these cities to the number of articles in which they were mentioned in the last 12 years of the Chicago Tribune, and the percentage of 67 University of Chicago students who recognized them (also in 1994). Table 1 shows the 12 largest cities in Germany (in order of 1994 population) taken from the complete data. Computing Spearman correlations over all 83 cities, the surrogate correlation was .79, the ecological correlation was .70, and the recognition correlation was .60. We observe that the Chicago residents' recognition memory is slightly more in tune with how frequently cities are mentioned in media than with the actual populations. This is to be expected since the newspapers, and not city populations, are a part of our observable everyday environment. Given that there are solid relationships between recognition memory and socially-interesting criteria, how might the mind exploit this fact to make inferences?

Table 1: The largest German cities ranked by population, along with the number of articles in which they appeared in the newspaper, and the percentage of people who recognized them.

City	Articles	Recognition (%)
Berlin	3484	99
Hamburg	1009	96
Munich	1240	100
Cologne	461	82
Frankfurt	1804	96
Essen	93	28
Dortmund	84	19
Stuttgart	632	63
Dusseldorf	381	81
Bremen	140	44
Duisburg	53	7
Hannover	260	88

The Recognition Heuristic

I turn now to a model of how people draw inferences from recognition memory. The task is inferring which subset of a class of objects scores highest on some criterion. For sets of two objects, this amounts to the two-alternative forced-choice paradigm. An example question would be: Which river is longer? A) The Nile B) The Isar. For questions of this sort, the *recognition heuristic* is simply stated:

If only one of the alternatives is recognized, then choose it.

Naturally, this heuristic is only sensible in domains where recognition is correlated with the criterion. If the correlation is negative (as it would be in the task of inferring which of two cities is smaller, for instance), then the unrecognized alternative should be chosen. The purpose of this article, however, is not to explore when people follow recognition, or if it is a conscious process, but rather to test if they follow it and to describe the mathematical consequences of doing so. Please note that my use of the term recognition differs from the standard usage in the psychological literature (which refers to identifying items from a previous session among distractors) and familiarity (which has to do with degrees of exposure). Here, let us focus on mere recognition, that all-or-none state which accompanies objects which have been experienced before. This sense of recognition seems to constitute a separate memory system, and its presence can be felt everyday life when one hesitates in recalling the identity of a recognized face. (The Scottish verb "to tartle", interestingly, refers to just this sort of hesitation.) The recognition heuristic, which rests on this simple and subjective sort of memory, seems like a bias or overly simple strategy. Is it?

Accuracy

The accuracy attainable from the recognition heuristic depends on two variables: how often the heuristic can be applied, and the degree of correlation between recognition and the criterion of interest. As a thought experiment, consider the case of an American who learns, one by one, to recognize the 100 largest cities in France. Every time she learns a new city, she is given a test consisting of all possible pairs of cities drawn from the 100 largest, and her task is to pick the larger city in each pair. Before she has learned any of the cities, she will have to guess on every pair, and thus score 50% correct, as represented by the leftmost point in Figure 2.

At a later time, represented by the middle point, she has learned the names of 50 of the 100 cities, and the cities she recognizes are larger than the cities she does not recognize in 90% of all possible pairs. Assume further that when she recognizes both cities in a pair, she is able to use further knowledge to pick the larger one 60% of the time.
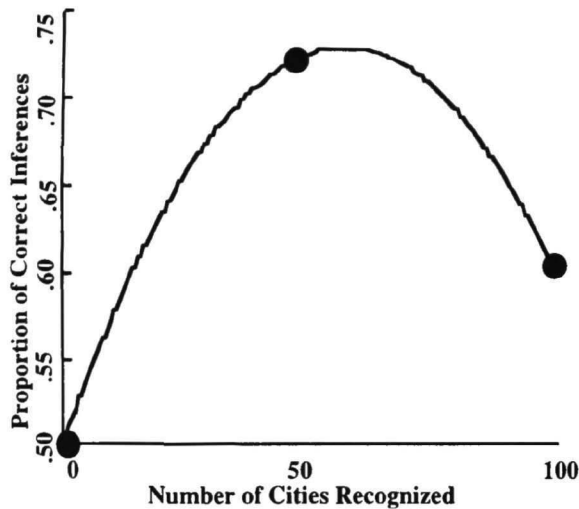
408

Figure 2: Applying the recognition heuristic may lead considerable accuracy, as well as a less-is-more effect.

A quick calculation shows that in half of the pairs on the quiz, she will recognize one city and not the other. She will get 90% of these pairs correct. In another one-quarter of the questions, she will recognize neither city, guess, and score 50% correct. On the remaining one-quarter, she will recognize both cities in the pair, and score 60% correct. At this intermediate state of recognizing half of the 100 cities, she will attain $(.5)(.9) + (.25)(.5) + (.25)(.6) = .725$, a respectable 72.5% correct. When she recognizes all 100 cities, represented by the rightmost point, she will recognize both cities in each pair, and score only 60% correct. The striking result is that she scored a higher percentage of accurate inferences when she recognized half the cities than she did when she recognized them all. Any state of affairs where lesser recognition knowledge enables more accurate inferences than greater recognition knowledge is a case of the *less-is-more effect*, which I shall try to evoke empirically. Before doing so, we must ask if the recognition heuristic is a fundamental mechanism in human inference.

A Simple Test of the Recognition Heuristic

This simple test asks how often unprompted people will use the recognition heuristic (Goldstein, 1996). I quizzed people on all pairs of cities drawn from the 25 (n=6) or 30 (n=16) largest in Germany (300 or 435 questions) and asked them to choose the more populous city in each case. Either before or after the test, the participants were asked to check off from a list which of these cities they recognized (order, however, had no effect). From this recognition information, I calculated how often participants had an opportunity to choose in accordance with the recognition heuristic, and compared it to how often they actually did. Figure 3 shows the results for 22 individual participants.

For each participant, two bars are shown. The lighter bar shows how many opportunities the person had to apply the recognition heuristic, and the darker bar shows how often their inferences agreed with the heuristic. For example, the

person represented by the leftmost pair of bars had 156 opportunities to choose according to the recognition heuristic, and did so every time. The next person did so 216 out of 221 times, and so on. The proportions of *recognition heuristic adherence ranged between 100% and 73%*. The median proportion of inferences following the recognition heuristic was 93% (mean 90%). Unprompted participants made most of their inferences in accordance with the recognition heuristic, perhaps for lack of a better strategy. Would they still follow it if predictive information which suggested violating recognition were available?
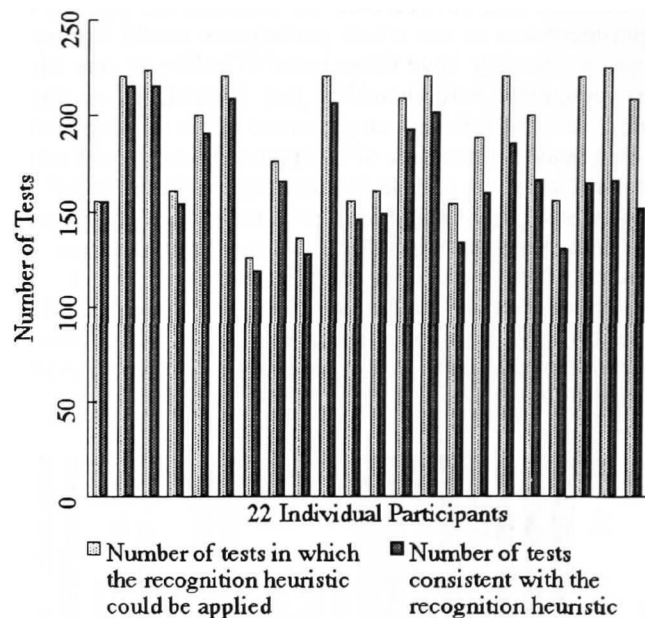

Figure 3: How often participants made choices in accordance with the recognition principle.

A Tougher Test of the Recognition Heuristic

In this experiment, participants were taught useful information that offered an alternative to following the recognition heuristic (Goldstein, 1996). The information was about the presence of major league soccer teams, which are powerful predictors of city population in Germany. The objective was to see which people would choose as larger: an unrecognized city, or a recognized city that they just learned has no soccer team.

The experiment began with a training session during which participants were instructed to write down all information that would follow. They were first told that they would be quizzed on the populations of the 30 largest cities in Germany. Next they were taught i) that nine of the 30 largest cities in Germany have soccer teams, ii) that the cities with teams are larger than cities without teams in 78% of all possible pairs, and iii) the names of four cities that have soccer teams, as well as the names of four cities that do not. These eight cities appeared to be randomly-drawn, but

409

the drawing was rigged so that the same eight, well-known cities were chosen by each participant. A pre-test followed in which participants had to reproduce all of the presented information exactly and could not proceed with the experiment until they did so. Either before or after the main task, participants were shown a list of German cities and asked to mark those that they recognized before coming to the experiment.

With their notes beside them, participants were then presented pairs of cities and asked to choose the larger city in each pair. To motivate them to take the task seriously, they were offered a chance of winning 15 dollars if they scored more than 80% correct. To reiterate, the point of the experiment was to see which participants would choose as larger: a city they have never heard of before, or one which they recognized beforehand but just learned had no soccer team. From the information presented in the training session (which made no mention of recognition), one would expect the participants to choose the unrecognized city. Why? An unrecognized city either has or does not have a soccer team. If it does, then there is a 78% chance that it is larger. If it does not, then there is an equal chance it is larger. If a person does not use recognition at all, the unrecognized city should always be chosen since any chance of it having a soccer team suggests that it is probably larger. Figure 4 shows the results.
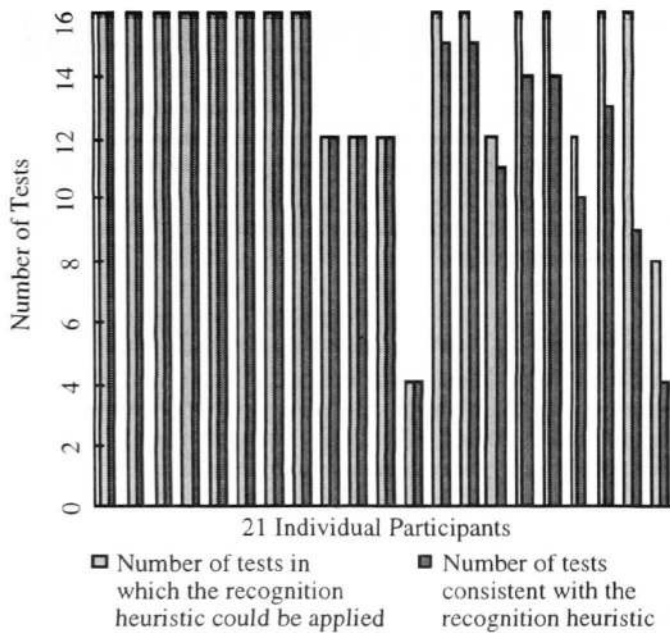

Figure 4: Recognition heuristic adherence despite training that encouraged the use of information conflicting with recognition.

In Figure 4, the left-hand bars are of different heights because individual participants recognized different cities before the experiment, so the number of cases where the recognition heuristic applied varied. Twelve of 21

participants made choices in accordance with the recognition heuristic without exception, most others deviated on only one or two items. All in all, participants chose in accordance with recognition in 273 of the 296 total critical pairs. The median proportion of inferences agreeing with the heuristic was 100% (mean 92%), despite conflicting knowledge. It appears that the additional information was not integrated into the inferences, consistent with the recognition heuristic.

Does the Less-Is-More Effect Occur in Human Reasoning?

As the previous studies show, the recognition heuristic can be a major force in certain inference tasks. This result provides empirical support to the theoretical prediction that the less-is-more effect should appear in certain situations. However, this effect is yet to be seen in the reasoning of people. Gerd Gigerenzer and I had 52 University of Chicago students take two quizzes each (Goldstein & Gigerenzer, in press). One was on the 22 largest cities in the US, cites about which they knew numerous facts useful for inferring population. The other was on the 22 largest cities in Germany, about which they knew little or nothing beyond name recognition. In fact, they did not even recognize about half of them. As in the previous two studies, each question consisted of two randomly-drawn cities, and the task was to pick the larger. If participants would score higher on the foreign cities (of which they recognized half) than on the domestic ones (of which they recognized all), it would be an instance of a less-is-more effect. The curious phenomenon of a less-is-more effect is hard to demonstrate with people who have definite knowledge of the criterion. That is, many Americans, and nearly all University of Chicago students, can name the three largest American cities in order. This alone gives them the correct answer for 26% of all possible questions. Those who know the top five American cities in order will score at least 41% correct from that knowledge alone. This definite knowledge of the criterion, coupled with the lifetime of knowledge Americans have about their own cities, would make their scores on the domestic test hard to match.

The result was that the Americans scored a median 71% (mean 71.1%) correct on the their own cities and a median 73% (mean 71.4%) correct on the foreign ones. Despite the presence of substantial knowledge about American cities, the recognition heuristic resulted in a very slight less-is-more effect. For half of the subjects, we kept track of which cities participants recognized: the mean proportion of inferences according with the recognition heuristic was 88.5% (median 90.5%). Furthermore, participants could apply the recognition heuristic nearly as often as possible because they recognized a mean of 12 cities, roughly half of the total (compare to Figure 2). In a study that is somewhat the reverse of this one, a less-is-more effect was demonstrated with Austrian students who scored a greater proportion of correct inferences on American cities than on German ones (Hoffrage, 1995).

The Recognition Heuristic As a Prototype of Fast and Frugal Heuristics

The recognition heuristic is one of many fast and frugal heuristics which organisms can use under limited time, knowledge, and computational might (Gigerenzer & Goldstein, 1996). Since it uses recognition memory, a fundamental psychological mechanism, and profits from missing knowledge, the recognition heuristic is perhaps the simplest of these adaptive tools. Reasoning by recognition is a form of one-reason decision making (Goldstein, 1996) because it bases complex inferences on recognition alone. This conflict-avoiding strategy eliminates the need to make trade-offs between cues pointing in opposing directions, a well documented desire of human decision makers (e.g., Baron, 1990; Hogarth, 1987; Payne, Bettman, & Johnson, 1993). In the interaction between minds and real-world environments, patterns of missing knowledge carry important information which organisms can exploit to make inferences. As this missing knowledge is filled in, the usefulness of recognition is diluted, and the accuracy and efficiency of inferences may decline.

References

Barnett, S. A. (1963). *The rat: A study in behavior.* Chicago: Aldine.

Baron, J. (1990). *Thinking and deciding.* Cambridge, UK: Cambridge University Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193-217.

Galef, B. G., Jr. (1987). Social influences on the identification of toxic foods by Norway rats. *Animal Learning & Behavior, 18,* 199-205.

Galef, B. G., Jr., McQuoid, L. M., & Whiskin, E. E. (1990). Further evidence that Norway rats do not socially transmit learned aversions to toxic baits. *Animal Learning & Behavior, 18,* 199-205.

Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 104,* 650-669.

Goldstein, D. G. (1996). Models of bounded rationality for inference. Doctoral dissertation, Department of Psychology, The University of Chicago.

Goldstein, D. G. & Gigerenzer, G. (in press). Recognition: How to exploit a lack of knowledge. In G. Gigerenzer & P. Todd (Eds.), *Ecological rationality: Simple heuristics that make us smart.* New York: Oxford University Press.

Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision.* 2nd ed. Chichester: Wiley.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* Cambridge: Cambridge University Press.