



Intuitive judgments of social statistics: How exhaustive does sampling need to be? ☆



Thorsten Pachur^{a,*}, Ralph Hertwig^a, Jörg Rieskamp^b

^a Center for Adaptive Rationality (ARC), Max Planck Institute for Human Development, Berlin, Germany

^b Center for Economic Psychology, Department of Psychology, University of Basel, Switzerland

HIGHLIGHTS

- We examine the role of noncompensatory processing in instance-based inference.
- We propose a heuristic for judging the relative frequency of events.
- The heuristic performs competitively in describing people's frequency judgments.
- In computer simulations we examine the ecological rationality of the heuristic.
- The heuristic performs competitively in clustered and skewed environments.

ARTICLE INFO

Article history:

Received 25 June 2012

Revised 1 July 2013

Available online 13 July 2013

Keywords:

Sampling

Frequency

Decision strategy

Heuristics

Availability

ABSTRACT

One way to make inferences about social statistics, such as the frequencies of health risks in the population, is to probe relevant instances in one's social network. People can infer, for instance, the relative frequency of different diseases by probing how many members of their social network suffer from them. How are such instance-based inferences cognitively implemented? Noncompensatory strategies based on lexicographic and limited search have been extensively examined in the context of cue-based inference. Their role in instance-based inference, by contrast, has received scant attention. We propose the *social-circle heuristic* as a model of noncompensatory instance-based inference entailing lexicographic and limited search, and test its descriptive and prescriptive implications: To what extent do people rely on the social-circle heuristic? How accurate is the noncompensatory heuristic relative to a compensatory strategy when inferring event frequencies? Two empirical studies show that the heuristic accurately predicts the judgments of a substantial portion of participants. A response time analysis also supports the assumption of lexicographic search: The earlier the heuristic predicted search to be terminated, the faster participants classified as using the social-circle heuristic responded. Using computer simulations to systematically investigate the heuristic's prescriptive implications, we find that despite its limited search, the heuristic can approximate the accuracy of a compensatory strategy in skewed and in spatially clustered environments—both common properties of distributions in real-world social environments.

© 2013 Elsevier Inc. All rights reserved.

Introduction

In myriad domains of social life, people's decisions are influenced by their observations of others. In fact, imitating the behavior of others

is a powerful and versatile heuristic that helps us to navigate the trials and tribulations of complex social environments (e.g., Hertwig & Herzog, 2009; Hertwig, Hoffrage, & the ABC Research Group, 2013; Richerson & Boyd, 2005). Knowing what others do, want, like, or have can help us make decisions in the face of such diverse issues as whether or not to adopt “green” behavior (Goldstein, Cialdini, & Griskevicius, 2008), whether to engage in helping behavior (Fischer et al., 2011), which cultural products (e.g., books, movies, TV shows, and music) to purchase and consume, and how satisfied we are with our income (Boyce, Brown, & Moore, 2010). In Salganik, Dodds, and Watts's (2006) investigation of simulated cultural markets, for instance, individuals' music preferences were substantially altered when given frequency information about the choices of other individuals in the market.

☆ We thank Florian Steinmann for collecting the data in Study 1, Henrik Olsson and Lael Schooler for their many constructive comments, and Laura Wiles, Valerie Chase, and Susannah Goss for editing the manuscript. This research was supported by a grant from the German Research Foundation (DFG) as part of the priority program “New Frameworks of Rationality” (SPP 1516) to Ralph Hertwig and Thorsten Pachur (HE 2768/7-1).

* Corresponding author at: Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

E-mail address: pachur@mpib-berlin.mpg.de (T. Pachur).

Sometimes we have the benefit of explicit information about the frequency of others' behaviors (e.g., how many people have visited a website or seen a movie on the opening weekend). In many real-life situations, however, we have no such objective social statistics at hand, and thus need to rely on much more limited counts of experiences stored in memory. In May 2011, for instance, German consumers may have wondered whether to stop consuming raw tomatoes, fresh cucumbers, and leafy salads, as recommended by the German Federal Institute of Risk Assessment after a sudden increase in life-threatening infections caused by Shiga-toxin-producing *Escherichia coli* (STEC; ECDC, 2011). Unable to look up an official count of others' choices in the same situation, consumers could gauge the frequency of recommendation-compliant behaviors among the members of their close social network (e.g., friends and family).

Compensatory and noncompensatory processing

Accessing frequency information in terms of instances experienced by one's proximate social network has been proposed as a key mental tool for inferring the frequency of behaviors or characteristics in the population (Fiedler & Juslin, 2005; Galesic, Olsson, & Rieskamp, 2012; Tversky & Kahneman, 1973). It is implicitly or explicitly assumed that this instance knowledge is processed in a *compensatory* fashion. Compensatory strategies consider all available information and process it such that conflicting pieces of evidence can be traded off against each other. Undeniably, compensatory strategies have been successful in describing people's frequency judgments (e.g., Hertwig, Pachur, & Kurzenhäuser, 2005; Pachur, Hertwig, & Steinmann, 2012). However, also *noncompensatory* strategies play a major role in human judgment and decision making (Ford, Schmitt, Schechtman, Hults, & Doherty, 1989; Gigerenzer, Hertwig, & Pachur, 2011). Noncompensatory strategies ignore parts of the information, with the consequence that a piece of evidence that supports one option cannot be compensated for by another piece that favors the other option but is ignored (e.g., Katsikopoulos, Pachur, Machery, & Wallin, 2008). Due to their limited search, noncompensatory strategies respect the boundaries of human information processing (Gigerenzer, Todd, & the ABC Research Group, 1999; Simon, 1990).

To date, the comparison of compensatory and noncompensatory strategies has been limited to cue-based inference¹ (Gigerenzer et al., 1999; Rieskamp & Hoffrage, 2008; see also Payne, Bettman, & Johnson, 1993; Tversky, 1972). Yet evidence for the use of noncompensatory strategies in cue-based inference (Einhorn, 1970; Gigerenzer et al., 2011) raises the question of whether and to what extent noncompensatory processing also occurs in the context of instance-based inference.

Limited search and noncompensatory processing may play a role in instance-based inference for several reasons. First, it can reduce processing cost. Indeed, noncompensatory processes are particularly evident in decisions involving information cost, such as cue-based inference from memory (e.g., Bröder & Schiffer, 2003, 2006); instance-based inferences are often memory-based. Second, due to information redundancy in natural environments, limited search can result in decisions that coincide with those based on more extensive search. As shown by Gigerenzer and Goldstein (1996) in the context of cue-based inference, structural aspects of the environment (e.g., intercorrelations between cues) permit simple mechanisms to approximate the accuracy of more complex strategies. Similarly, in a risky choice context, Hertwig and Pleskac (2008, 2010) found that increases in inferential accuracy level off relatively quickly with increasing sample sizes. This decreasing marginal utility of more information

¹ In contrast to instance-based inference, cue-based inference relies on semantic properties of an event or object to make an inference. For instance, to judge whether there are more people in Germany who belong to a basketball or a tennis club, a cue-based strategy would consider properties of the respective sports (e.g., whether it is a team sport or an individual sport) as cues.

may also hold for instance-based inferences of event frequencies. Moreover, Hertwig and Pleskac (2010) demonstrated that small samples facilitate decision making because they permit decision makers to discriminate among options more easily. A final advantage of noncompensatory processing of instance-based knowledge is reliability of knowledge. Specifically, sample spaces may differ in terms of how reliable a decision maker's knowledge of them is. Ordered and noncompensatory processing of these sample spaces elegantly enables the decision maker to prioritize sample spaces with more reliable knowledge.

Our goals in this article are the following. First, we propose a model of a heuristic that represents noncompensatory processing of instances: the *social-circle heuristic*. Second, we investigate to what extent this heuristic is a descriptively accurate model of people's judgments of social statistics (frequencies). Third, from a prescriptive perspective, we determine the price (in terms of accuracy) the heuristic pays for ignoring part of the information. To these ends, we examine how well the social-circle heuristic fares in accounting for people's inferences, relative to a compensatory instance-based heuristic (Studies 1 and 2) and to three cue-based strategies (Study 2). In Study 3, we turn to a systematic analysis of the prescriptive question by addressing the *ecological rationality* of the social-circle heuristic (Todd, Gigerenzer, & the ABC Research Group, 2012). Specifically, we use computer simulations to investigate the environmental structures fostering and hampering the heuristic's performance, relative to that of a compensatory strategy. We begin by describing the established compensatory models of instance-based inference.

Models of compensatory processing of instance-based inference

Perhaps the most prominent account of how people infer the frequency of a class of events (or the probability of an event) is Tversky and Kahneman's (1973) *availability heuristic*. It assumes that when judging the frequency with which an event category occurs in the world (e.g., heart attacks among middle-aged people), people recall the event's occurrences from memory. The likelihood that an occurrence (instance) stored in memory is sampled is a function of its "availability," and this ease of retrieval depends on, for instance, how "vivid" or "dramatic" the instance is. In principle, however, the mnemonic sample space is limited only by the bounds of the decision maker's knowledge and can include directly experienced as well as "virtual" instances (e.g., media reports about a person killed in a shark attack).

Another way to model instance-based frequency judgments is in terms of exemplar models (e.g., Nosofsky, 1986; Pachur & Olsson, 2012). In these models, memory representations contribute to judgments of frequency as a function of their similarity to the class of target events. For instance, Dougherty, Gettys, and Ogden's (1999) MINERVA-DM model assumes that each individual encounter with an instance—regardless of whether the instance is real or virtual—is stored and that the target event category (e.g., heart attack) is compared with these memory traces using a global matching process. Event categories that result in a stronger activation, or "echo," in memory are inferred to be more frequent. Dougherty et al. (1999) showed that phenomena that are usually explained in terms of the availability heuristic can also be accounted for by MINERVA-DM.² Like the availability heuristic, exemplar models assume that the sampling space is restricted only by the bounds of a person's knowledge.

² According to another exemplar model, Juslin and Persson's (2002) PROBEX, inferences about the population frequency of event categories can be made by retrieving knowledge about other, similar event categories. For instance, in order to infer whether tuberculosis or bladder cancer occurs more frequently, PROBEX would retrieve criterion knowledge about the frequencies of other infectious diseases and cancers and integrate this knowledge as a function of their similarity to tuberculosis and bladder cancer, respectively. Because instance knowledge does not enter the processing directly, we will not treat PROBEX as a genuine instance-based model.

Hertwig et al. (2005) proposed a compensatory strategy that assumed a more constrained sample space. *Availability-by-recall*, a specific instantiation of the availability heuristic (see also Sedlmeier, Hertwig, & Gigerenzer, 1998), restricts the sampling space to directly experienced instances and excludes virtual ones.³ To infer, say, whether more people in the general population suffer from tuberculosis or bladder cancer, availability-by-recall searches for instances of both diseases within a person's proximate social environment—defined as his or her family, friends, and acquaintances. Disease instances beyond this close network, such as those involving strangers or those featured in the media, will not be sampled. Hertwig et al. compared availability-by-recall with three alternative cognitive strategies and found that, side-by-side with another strategy (which assumed an automatic processing of actual frequencies), it provided the best description of people's judgments of risk frequencies. Relatedly, Pachur et al. (2012) found that a strategy that—like the availability heuristic—additionally recruits instance knowledge from the media (e.g., news, movies, internet) was much less predictive of people's judgments than was availability-by-recall. Therefore, we will not consider the availability heuristic here.

Like the availability heuristic and exemplar models, availability-by-recall entails compensatory processing: it sums up the relevant instances in the defined sample space (i.e., a person's social network). As a consequence, the non-occurrence of instances of, say, bladder cancer in a person's close family can be compensated for by a case of bladder cancer among his or her friends. Despite the empirical support for the compensatory mechanism of availability-by-recall in Hertwig et al. (2005), there are indicators that (at least some) people process instances in a noncompensatory fashion. Take, for instance, the *false consensus effect*. In the classic demonstration by Ross, Greene, and House (1977), students estimated the percentage of other students who agreed with a statement as being higher (lower) when they themselves agreed (vs. disagreed) with it (63.5% vs. 23.3%; for an overview, see Marks & Miller, 1987). Importantly, this pattern also emerged when respondents were informed about others' opinion regarding the statement (e.g., Krueger & Clement, 1994, Experiment 3), that is, when they enjoyed explicit instance knowledge. One interpretation of these results is that some people may rely on a very small sample of instances (a sample that first and foremost includes themselves) rather than considering their entire store of instance knowledge. Assuming that judgments of event frequencies sometimes stem from limited search within a person's mnemonic repertoire of instances, how can such limited search and noncompensatory processing of instances be modeled? In the following, we propose one possible model.⁴

Noncompensatory processing of instances: the social-circle heuristic

The social-circle heuristic aims to model how people judge the relative frequency of others' beliefs, preferences, behaviors, and characteristics in

³ In light of evidence for mental contamination (e.g., Wilson & Brekke, 1994), one may ask how this targeted sampling process is implemented. One possibility is that search in memory is restricted by the use of contextual cues (e.g., family, friends). Through those cues, relevant instances outside the intended sampling space (e.g., those encountered in the media) receive little or below-threshold activation (e.g., Dougherty et al., 1999). In general, people making frequency judgments seem well able to discount information from sources deemed nonrepresentative (e.g., Oppenheimer, 2004), indicating that contextual information shapes the search process.

⁴ It is worth pointing out that exemplar models such as MINERVA-DM could in principle implement a noncompensatory processing of instances by attaching noncompensatory weights to relevant instances (cf. Martignon & Hoffrage, 2002). However, for the sake of conceptual clarity and because no exemplar model has yet been proposed that weights relevant instances in such a way, we do not include exemplar models in the class of *noncompensatory* instance-based strategies. For a discussion of exemplar processing based on limited search in categorization, see De Schryver, Vandist, and Rosseel (2009) and Juslin and Persson (2002).

the population. Specifically, it predicts which of two event categories is inferred to occur more frequently in a reference class. Fig. 1 depicts the heuristic's processing steps. In order for the heuristic to be applicable, the names of both event categories need to be recognized. If only one category is recognized, the inference is assumed to be based on the recognition heuristic (see Pachur, Todd, Gigerenzer, Schooler, & Goldstein, 2011); if none is recognized, an inference is made by guessing. Like availability-by-recall, the social-circle heuristic searches for instances of both events within a person's proximate social environment. Unlike availability-by-recall, however, the heuristic entails lexicographic and limited search. Specifically, it exploits the well-documented hierarchical structure of social networks, which consist of discrete subgroups ("circles") of increasing size (e.g., Hill & Dunbar, 2003; Milardo, 1992; Zhou, Sornette, Hill, & Dunbar, 2005)—in line with the increasing evidence of a link between the structure of the external world and search in memory (e.g., Hills, Todd, & Goldstone, 2008; Pirolli & Card, 1999). These circles are sequentially probed for critical instances. Search *within* a circle is not limited, and all relevant instances in a circle enter the final tally for this circle. Search *across* circles can be limited (this appears plausible because it has been found that people employ social categories to probe social memory; Bond & Brockett, 1987; Fiske, 1995; Hills & Pachur, 2012).

What are these circles? Consistent with the observation that people consider their own behaviors when inferring the prevalence of behaviors at large (Ross et al., 1977), the social-circle heuristic begins by probing information about the self (circle 1). From this starting point, two alternative social dimensions may guide further search: (a) altruism or (b) frequency of contact. Altruism typically manifests itself in kin relationships (Hamilton, 1964) and in reciprocal relationships (Singer, 1981). Using the dimension of altruism to describe social network structure, one can hypothesize three circles beyond the "self" circle: family (circle 2), friends (circle 3), and acquaintances (circle 4), with family encompassing kin relationships (including nongenetic relationships, that is, with one's partner and his or her family) and friends and acquaintances encompassing nonkin reciprocal relationships (see Geary & Bjorklund, 2000; Kahn & Antonucci, 1980; Moreno, 1936). Based on this structure, our first instantiation of the social-circle heuristic implements search in terms of a hierarchy of altruism-based circles (Fig. 1). One advantage of this lexicographic order is that people are likely to have the most reliable, extensive, and easily retrievable instance knowledge about themselves, followed by family members, close friends, and acquaintances, respectively (Henrich & Henrich, 2007, p. 58).

Our second implementation of the social-circle heuristic assumes that a person's social network is partitioned according to frequency of contact. This structure acknowledges that the individuals about whom a person has the most extensive knowledge need not be family. They could, for instance, be close friends or colleagues. Using frequency of contact as the dimension structuring social networks, the *social-circle heuristic_F* (with *F* for frequency) assumes, beyond the self circle, the circles of people with whom one has contact at least once a week (circle 2), about once a month (circle 3), and no more than once in six months (circle 4). The rationale of this lexicographic order is that frequency of contact has been demonstrated to be a key determinant of the retrieval probability of memory records (Anderson & Lebiere, 1998; Anderson et al., 2004).

Studying search in social memory, Hills and Pachur (2012) found evidence suggesting that both search orders (altruism and contact frequency) are consistent with how people sequentially retrieve the individuals they know. Specifically, in a free recall task, participants tended to name their family members earlier than their friends and acquaintances, but they also named people earlier, the more frequent their contact with them.

We now turn to the architecture of the noncompensatory social-circle heuristic. For illustration, consider the following inference task: "In Berlin, do more people drink (a) Coca-Cola or (b) Pepsi?" This task

Which event category is more frequent in the population: A or B?

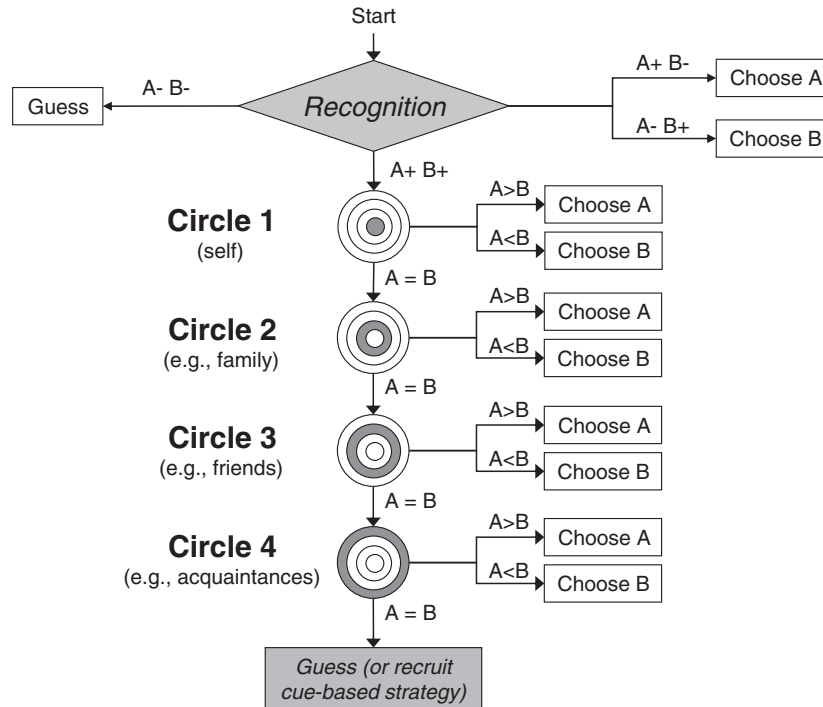


Fig. 1. Flow chart showing the social-circle heuristic and the relationship of its sampling process to the recognition principle and inferences based on other cues.

is an example of the general problem of inferring which preferences, behaviors, or social events are more frequent in a population. Let us assume that a person has heard of both brands and so cannot apply the recognition heuristic. According to the social-circle heuristic, this person will search her instance knowledge (e.g., consumers of Coca-Cola vs. Pepsi in her social network) and use the respective tallies within each circle to select alternative *a* or *b*. The architecture of both instantiations of the social-circle heuristic (based on altruism vs. contact frequency) can be described by the following three building blocks: a search rule, a stopping rule, and a decision rule. Table 1 describes these building blocks, which specify how search proceeds, how search is terminated, and how a final decision is made.

Due to its stopping rule, the heuristic embodies limited search. Specifically, search is stopped as soon as within a circle the tally (i.e., the number of instances) for one alternative differs from the tally for the other alternative, irrespective of the size of the difference. The tallies for the individual circles are not integrated across circles. Therefore, unlike (fully) compensatory strategies such as availability-by-recall (which integrates instances across all circles), the social-circle heuristic is a noncompensatory strategy (although it is locally compensatory in its exhaustive within-circle search). Instances accessible in later circles cannot reverse an inference based on earlier circles. Thus, the social-circle heuristic and availability-by-recall recruit instances from the same sample space—individuals in a person's proximate social network—but, unlike availability-by-recall, the social-circle heuristic may leave some social circles uninspected.

Next, we test the social-circle heuristic in terms of (a) how well it describes people's inferences, and (b) how well it approximates the accuracy achieved by the compensatory availability-by-recall strategy. In Study 1, we investigate the descriptive power of the social-circle heuristic relative to availability-by-recall; in Study 2, we extend the model comparison to the social-circle heuristic_F and three cue-based strategies.

Study 1: compensatory versus noncompensatory processing of instances in inferences about cancer mortality

Participants were asked to judge the relative mortality rates of different types of cancer, a domain in which Hertwig et al. (2005) have found support for compensatory availability-by-recall (see also Pachur et al., 2012). We also probed participants' instance knowledge of cases of death from cancer in their social networks. As the target criterion in the inference task refers to deaths resulting from cancer, the first circle (i.e., the self circle) of the social-circle heuristic never discriminated here. Importantly, in constructing the test bed for examining the strategies, we aimed to implement a representative design (cf. Brunswik, 1955; Dhimi, Hertwig, & Hoffrage, 2004) encompassing a broad set of events instead of focusing on a small and selective one. To that end, we employed all 24 types of cancer tracked in the Swiss national statistics (e.g., Bundesamt für Statistik, 2004; see Hertwig et al., 2005). By using this full set of cancers, we were also able to

Table 1
The three building blocks of the social-circle heuristic.

Search rule	Search the social circles for occurrences of the target events (e.g., people who drink Coca-Cola and Pepsi, respectively), proceeding sequentially from the first circle ("self") through the other circles. Within each circle, tally the number of instances of each of the two target events.
Stopping rule	If the tally of relevant instances of the two target events is unequal within the first circle (e.g., I drink only Coca-Cola), then stop search and go on to the decision rule. Otherwise, search subsequent circles one by one, and proceed to the decision rule as soon as the tally of instances of the two events is unequal within a circle. If circle 4 does not yield unequal tallies, proceed to the decision rule.
Decision rule	Infer that the event category with the higher tally is the more frequent one in the reference class. If the tallies of sampled instances do not differ in circle 4, then guess (or proceed to cue knowledge; see Appendix C).

gauge how accurately the instance-based strategies discriminate across a wide range of real-world risks and how often they are applicable.

Method

Participants

Thirty-three students participated. The study was conducted at the University of Basel, Switzerland. In addition to course credits, participants received performance-contingent payment, earning 0.04 Swiss francs (= US\$ 0.05) for every correct inference and losing the same amount for every incorrect inference.

Materials

Table 2 shows the 24 types of cancer and their respective annual frequencies of death in Switzerland, averaged across six consecutive years (1999–2004) to smooth out year-to-year fluctuations. We constructed a set of all possible 276 pairs of cancers. Participants completed two different tasks. In the *inference task*, they were asked to infer which of two cancers has a higher annual mortality rate in Switzerland. In the *retrieval task*, they indicated, for each type of cancer and for each of three social circles (family, friends, and acquaintances), how many people (if any) they knew who had died from the respective cancer (instance knowledge).

Procedure

After reading an introductory text explaining the relevance of accurate judgments of the relative frequencies of different cancers, participants were instructed as follows:

“We ask you to judge the annual mortality rate of different types of cancers in Switzerland. ... Each item lists two different types of cancer. The question you are to answer is: For which of the two cancers is the number of deaths per year higher?”

Pairs of cancer types were displayed sequentially on a computer screen in 12 blocks of 23 pairs. After completing each block, participants took a short break (without leaving the laboratory). They indicated their judgment by pressing one of two designated keys on the keyboard.

Table 2

The 24 types of cancer used in Study 1, their respective frequencies of death in Switzerland (averaged across the years 1999–2004; e.g., Bundesamt für Statistik, 2004), and the number of instances participants recalled from their own social networks.

Type of cancer	Annual mortality rate	Number of recalled instances
Lung cancer	2756.0	17
Breast cancer	1347.3	29
Leukemia and lymphoma	1331.7	13
Prostate cancer	1312.3	11
Colon cancer	1172.2	3
Pancreatic cancer	897.8	9
Stomach cancer	572.2	6
Liver cancer	513	6
Cancer of the central nervous system	455	11
Ovarian cancer	453.2	1
Bladder cancer	450.5	1
Rectal cancer	437.2	0
Esophageal cancer	384.5	3
Cancer of the mouth and throat	351	0
Renal cancer	339.2	1
Cervical cancer	295.8	5
Skin cancer	242	8
Gall bladder cancer	196.5	3
Cancer of the connective tissue	94.3	0
Laryngeal cancer	94.2	8
Thyroid cancer	69.3	1
Bone cancer	37.5	4
Testicular cancer	17.2	1
Penile cancer	10.2	0

The order in which the cancer types appeared within a pair was determined at random for each participant, as was the order in which pairs were presented. The retrieval task was administered as a paper-and-pencil questionnaire. It included all 24 types of cancer and always followed the inference task. Sessions lasted about 60 min.

Results

On average, participants selected the cancer with the higher mortality rate in 62.2% ($SD = 4.5$) of cases. The resulting bonus was, on average, 7.68 Swiss francs (= US\$ 9.18). The modest level of accuracy indicates that participants most likely did not possess direct knowledge of the different cancers' mortality rates, thus providing a good test bed for comparing strategies that make inferences under conditions of limited knowledge.

Each participant reported, on average, 4.2 ($SD = 3.4$) members of their social network who had died from one of the 24 types of cancer. To analyze how good a proxy these retrieved samples were for judging which of two cancers has a higher mortality rate in the population, we correlated them (collapsed across participants; see Table 2) with the actual mortality rates. The number of reported instances and actual frequencies were strongly correlated ($r = .68$, $p = .001$; Spearman's rank correlation $r_s = .63$, $p = .001$).

How well do the strategies predict people's inferences?

We used each participant's reported instance knowledge to determine the predictions of the compensatory availability-by-recall and the noncompensatory social-circle heuristic. The two strategies made predictions in, on average, 25.2% and 26.1% of all 276 pair comparisons, respectively. (The strategies made no prediction when a person recalled no or the same number of instances for both cancers in a pair comparison.) Among those cases and separately for each participant, we computed the percentage of correctly predicted inferences. On average, the social-circle heuristic correctly predicted 74.5% ($SD = 15.8$) of people's inferences, slightly less than availability-by-recall, at 75.2% ($SD = 11.9$). One key reason for this nearly identical level of performance is that, although the social-circle heuristic relied on less information than did availability-by-recall (see below), they nearly always made the same prediction ($M = 99.5\%$, $SD = 1.1$). Consequently, when we classified each participant to the two strategies using a maximum likelihood approach (see Appendix A for details), availability-by-recall and the social-circle heuristic showed identical fit for 25 of the 29 participants who reported instance knowledge (86.2%). Three participants' (10.3%) inferences were best described by availability-by-recall, and one participant's (3.5%) by the social-circle heuristic. Four participants reported no instance knowledge and thus were not considered in this analysis.

How accurate and how frugal are the strategies' inferences?

Of all cases in which they made a prediction, the social-circle heuristic correctly predicted the more frequent cancer type in 71.4% of cases ($SD = 17.9$) and availability-by-recall in 71.9% ($SD = 18.0$) of cases. How often did the social-circle heuristic truncate search? To answer this question, we determined (based on the instance knowledge reported by each participant) the *stopping rate* for each circle, defined as the proportion of cases in which the social-circle heuristic stopped search at this circle (given it led to an unambiguous prediction). The heuristic terminated search at the second, third, and fourth circles in 43%, 12%, and 45%, respectively, of those cases in which it made a prediction. Thus, although the social-circle heuristic and availability-by-recall often made identical predictions, in 55% of cases the former inspected only part of the available instance knowledge. Due to its limited search, the social-circle heuristic considered (slightly) less information than availability-by-recall, with, on average, 1.13 ($SD = 0.23$) versus 1.17 ($SD = 0.29$) retrieved instances per inference, respectively, $t(33) = -2.14$, $p = .04$.

Table 3
Circle validities and discrimination rate (DR) for the social-circle heuristic in Study 1 and for the social-circle heuristic and its variant, the social-circle heuristic_F, in Study 2.

Circle	Study 1 (cancer mortality)		Study 2 (sports popularity)			
	Social-circle heuristic		Social-circle heuristic		Social-circle heuristic _F	
	Validity (SD)	DR (SD)	Validity (SD)	DR (SD)	Validity (SD)	DR (SD)
1	–	–	0.51 (0.19)	0.04 (0.05)	0.51 (0.19)	0.04 (0.05)
2	0.78 (0.18)	0.11 (0.12)	0.66 (0.24)	0.11 (0.11)	0.60 (0.20)	0.23 (0.16)
3	0.76 (0.13)	0.04 (0.07)	0.60 (0.18)	0.29 (0.18)	0.60 (0.22)	0.22 (0.13)
4	0.69 (0.20)	0.13 (0.15)	0.55 (0.17)	0.38 (0.21)	0.55 (0.16)	0.40 (0.21)

The similar levels of accuracy achieved by the social-circle heuristic and availability-by-recall hint at an important property of the environment: information redundancy, which enables a strategy that limits search to approximate the performance of a compensatory strategy (see also Dieckmann & Rieskamp, 2007). Sampling the entire social network provided little additional information for availability-by-recall. Table 3 shows the circle validity for each search space assumed by the social-circle heuristic. It is defined as the proportion of correct inferences when an inference is made based solely on the tallies within a given circle. Surprisingly, the circle validities decreased from circle 2 to circle 4. This suggests that limiting search when the retrieved tallies in a given circle discriminate (see Table 3 for discrimination rates) does not need to exact costs in terms of lower accuracy. One possible reason is that circles ranked high in the hierarchy (i.e., self and family) discriminated in those comparisons that included one very common event category; thus, the high-ranked circles siphoned off the “easy” comparisons (i.e., where the difference in objective frequency is large). We return to this possibility in Study 3.

In sum, we obtained three results: First, the noncompensatory social-circle heuristic described people's judgments of social statistics about as well as the compensatory availability-by-recall strategy. Second, despite lexicographic and limited search, the social-circle heuristic often made the same inference as availability-by-recall and thus achieved nearly the same level of accuracy. This suggests that—due to information redundancy in the environment—searching more is not inevitably better. Third, more proximate social circles offered higher validity than more peripheral social circles, even though the latter are likely to represent larger sample spaces than the former (e.g., a person's circle of friends is likely to be larger than her family circle; Hills & Pachur, 2012; Zhou et al., 2005).

In Study 1, we examined the social-circle heuristic in inferences regarding cancer mortality, a domain in which previous studies have found support for a compensatory instance-based strategy (Hertwig et al., 2005; Pachur et al., 2012). One limitation of this domain, however, is that respondents were often unable—luckily for them and for the people in their social networks—to retrieve multiple instances of the respective risks. In fact, for 17% of the events in the reference class, no instances at all were retrieved, resulting in a relatively low applicability for both availability-by-recall and the social-circle heuristic. This finding reveals that instances often cannot be recruited across a whole range of events. Although this is an important result, the low applicability of the strategies in the cancer domain may have limited our ability to contrast them.

Study 2: compensatory versus noncompensatory processing of instances in inferences about popularity of sports

We now investigate the social-circle heuristic in a context in which respondents are likely to have a richer repertoire of instance knowledge: the relative popularity of participative sports. Again we test the social-circle heuristic against the availability-by-recall strategy, but to the end of a more encompassing model comparison test we also test it against the social-circle heuristic_F (where search is guided by frequency of contact rather than an altruism structure) and against three cue-based strategies (Gigerenzer et al., 1999; see Appendix B for

a more detailed description of these strategies). Rather than drawing on known instances of the event categories, cue-based strategies use generic properties of the categories (e.g., whether a sport is a ball game) to infer their frequencies. The cue-based *take-the-best* heuristic is a noncompensatory strategy that searches for cues in the order of their perceived validity (i.e., the subjective probability that the cue leads to a correct response; Appendix B) and, when a discriminating cue is found, chooses the alternative with the positive cue value. The compensatory *tallying* heuristic considers all cues and tallies up the number of positive cue values, subtracts the number of negative values, and chooses the alternative with the higher sum. The compensatory *Franklin's rule* assumes that before the values of all cues are summed, they are multiplied by their respective validities; it chooses the alternative with the larger sum. Finally, we also investigate the social-circle heuristic's assumed lexicographic processing of instances by inspecting response times.

Method

Participants

Forty students of various disciplines in Berlin universities were recruited. The study was conducted at the Max Planck Institute for Human Development. In addition to a flat fee of €9 (=US\$ 11.10), all participants received the same performance-contingent payment as in Study 1.

Materials

Table 4 lists the 25 most popular participative sports in Germany. Popularity is defined in terms of the number of officially registered members of clubs for the respective sport. Information about sports club participation was taken from official statistics (e.g., Statistisches Bundesamt, 2002) and averaged across five consecutive years (1997–2001) to smooth out year-to-year fluctuations. In the *inference task*, participants were asked to judge, for each of the 300 possible pairs of sports, which of the two sports is more popular in Germany. In addition, we asked participants to complete three other tasks, each necessary to derive predictions for the models tested. In the *recognition task*, participants indicated whether they had previously heard of the sport.⁵ In the *retrieval task*, they indicated how many people (if any) in their social network were members of a club for each sport. Specifically, they reported for each of the four network circles (self, family, friends, and acquaintances) how many instances of sports club members they could recall. In order to instantiate the social-circle heuristic_F, we also asked participants to indicate for each recalled instance how often they typically had contact with that person. “Having contact” was defined as talking to the person for at least 5 min or sending to or receiving from the person a message of at least 100 words in length (see Pachur, Schooler, & Stevens, 2013). Frequency of contact was rated on a five-point scale with the categories “several times a week,” “once a week,” “approximately once a month,” “around once in six months,” and “less than once in six

⁵ In Study 1, we assumed that our student participants recognized the different types of cancer (but did not test this assumption).

Table 4

The 25 most popular participative sports in Germany, the number of active and passive club members (averaged across the years 1997–2002; e.g., *Statistisches Bundesamt, 2002*), and the number of club members participants recalled from their own social networks.

Sport	Number of club members	Number of recalled instances
Soccer	6,234,883	180
Gymnastics	4,800,199	12
Tennis	2,085,327	58
Shooting	1,584,931	10
Athletics	851,075	29
Handball	833,345	70
Equestrian	735,229	48
Table tennis	710,267	12
Skiing	677,556	18
Sports fishing	650,921	22
Water sports	633,652	90
Volleyball	530,399	34
Golf	320,630	41
Judo	268,475	49
Bowling	266,538	25
Dancing	255,190	57
Badminton	230,058	27
Basketball	202,938	93
Sailing	190,577	42
Ice sports	173,625	21
Cycling	153,141	27
Canoe	111,545	8
Karate	106,582	26
Chess	94,172	5
Rowing	78,746	21

Note: Eight participants indicated that they were members of a sports club themselves (three for water sports, two for basketball, two for judo, one each for soccer, sports fishing, athletics, and dancing; three participants reported being a member of two sports clubs).

months.” (Responses were later collapsed to conform to the circle definition of the social-circle heuristic_F.)

Finally, the *cue assessment task* consisted of four components. First, participants assessed each of the 25 sports on eight cues (cues were obtained in a supplementary study described shortly). Specifically, they assigned a (binary) cue value to each sport. For instance, they stated whether, in their opinion, soccer is a team sport or an individual sport. Second, they assessed the directionality of the cues, stating which of the two values on each cue (e.g., team sport or individual sport) is indicative of a higher number of club members. Third, participants rank-ordered the cues according to their perceived validity. Finally, they estimated the validity of each cue (“In how many of 100 pairs in which one sport has a positive cue value and the other a negative cue value does this cue make a correct prediction concerning the popularity of the sport?”).

We identified the cues that may be used to infer the popularity of different sports by surveying 30 students prior to Study 2. They were presented with a list of the 25 sports and asked to generate properties of the sports that they deemed to be indicative of which of two sports is more popular. The eight most frequently generated properties (cues) are reported in Table 5 and were used in the cue assessment task.

Procedure

The inference, recognition, and cue assessment tasks were presented on a computer screen, and the retrieval task was administered as a paper-and-pencil questionnaire. For the inference task, participants read the following instruction:

“We ask you to judge the popularity of various sports in Germany. ... Each item lists two different sports. The question you are to answer is: For which of the two sports is the number of club members in Germany higher?”

Three hundred pairs of sports were presented sequentially in 12 blocks of 25 pairs. After completing each block, participants could take a short break. The elements of each pair were presented on the left versus right side of a computer screen, with the order being determined randomly. Participants indicated each judgment by pressing one of two designated keys on the keyboard, and we recorded the time they took to make their inferences. In the recognition task, sports were presented sequentially in random order. The inference, recognition, retrieval, and cue assessment tasks were always administered in this fixed order.

Results

On average, respondents selected the sport with the higher number of club members in 62.9% ($SD = 6.3$) of cases and earned a bonus of, on average, €3.10 (= US\$ 3.8; $SD = 1.52$). This level of accuracy is similar to that observed in Study 1, and again provides a good test bed for comparing the strategies. With the exception of two sports (shooting and sports fishing), all participants recognized all sports; only pairs in which both sports were recognized (on average, 97.2% of the 300 pairs) were included in the analyses.

Table 5 reports the results for the cue assessment task. Respondents strongly agreed about the directionality of the cues, indicated by a Kendall's W of .59 ($p = .001$) calculated across all cues.⁶ When respondents were asked to rank-order cues according to their validity, no such consensus emerged. Kendall's W across participants was .02 ($p = .60$). We next calculated the validities of the eight cues (Table 5): Using each participant's stated cue values and cue directions, we determined the cues' validities and their discrimination rates. A cue's validity is the proportion of times in which a sport with a positive cue value is objectively more popular than a sport with a negative value on the same cue. A cue's discrimination rate is the relative frequency with which it discriminates between the two sports in a pair. The mean validities (across participants) ranged from .67 (ball games) to .49 (team sports), with an average validity of .59.⁷

How valid are participants' samples of instances relative to this cue knowledge? The numbers of club members that participants recalled from their social networks are reported in Table 4. On average, each participant recalled 25.6 relevant instances ($SD = 19.1$), about six times as many as in Study 1. To analyze how predictive these samples were for judging which of two sports is more popular, we compared the distribution of recalled instances across the different sports (collapsed across all participants) with the distribution in the population. The two were moderately correlated ($r = .53$, $p = .006$; $r_s = .26$, $p = .20$); the correlation was lower than in Study 1 ($r = .68$). Second, we calculated for each participant how often each individual circle would lead to an accurate inference and how often it would render possible a prediction. Table 3 shows the average circle validities and discrimination rates separately for the

⁶ For the model comparison reported below, cues were recoded (for each participant) such that a positive cue value pointed in the direction deemed to be indicative of a higher number of sports club members.

⁷ The validities appear relatively low. Note, however, that these subjective cue validities are likely to differ from ecological validities, which are based on objective cue values (Gigerenzer, Hoffrage, & Goldstein, 2008). We estimated the ecological validities of the cues using the modal response obtained in the cue assessment task. For all eight cues, the ecological validity was higher than the average cue validity. For instance, the ecological validities of the cues “national star,” “school sport,” and “seasonal dependency” were .76, .74, and .72, respectively. Thus, the low cue validities based on participants' cue knowledge were due to individual differences in the cue assessment task rather than to a genuine lack of predictive power in the cues. Did participants' estimates of the cues' validities match the calculated ecological validities? The answer clearly seems to be no. Across participants, we found no correlation between the estimated and the ecological cue validities (average correlation: $r = -.03$; $r_s = .07$). The mean estimated validities ranged from .72 (team sports) to .60 (outdoor sports), with an average validity of .65.

Table 5
Predictive directions, validities, and discrimination rates (DR) of the cues in Study 2. The cue directions indicate participants' modal response and the values in brackets are the proportions of participants giving that response.

Cue	Direction	Validity (SD)	DR (SD)	Estimated validity (Mdn)	Estimated rank (Mdn)
National star (Are there any famous German names in the sport?)	+ (0.93)	0.65 (0.12)	0.48 (0.04)	65	5
School sport (Is the sport offered at schools?)	+ (0.85)	0.66 (0.18)	0.45 (0.05)	64.5	5
Seasonal dependency (Is the sport season-dependent?)	– (0.98)	0.64 (0.10)	0.40 (0.10)	65	4
Ball game (Is the sport a ball game?)	+ (0.95)	0.67 (0.10)	0.41 (0.07)	65	5
Special equipment (Does the sport require special equipment?)	– (0.93)	0.57 (0.09)	0.49 (0.04)	65	4.5
Olympic sport (Is the sport an Olympic discipline?)	+ (0.80)	0.57 (0.13)	0.43 (0.08)	65	4
Outdoor sport (Is the sport played mostly outdoors or indoors?)	+ (0.60)	0.51 (0.09)	0.50 (0.03)	60	6
Team sport (Is the sport a team sport or an individual sport?)	+ (0.85)	0.49 (0.11)	0.40 (0.08)	72	3.5

two variants of the social-circle heuristic. The circle validities were considerably lower than in Study 1. In fact, circle 1 did not exceed chance level (.51). As in Study 1, the circle validities decreased from circle 2 (.66) to circle 4 (.55). These ecological analyses show that participants' personal samples of instances were less valid than their cue knowledge.

How well do the strategies predict peoples' inferences?

Table 6 shows that the three instance-based and three cue-based strategies predicted participants' inferences equally well (with the exception of take-the-best, which fared worse). The same pattern of results emerged when we included only those pairs in which (per participant) each of the six strategies arrived at a prediction (on average, 39% of all pairs; Table 6).

The strategies differed considerably with regard to how often they made a prediction. Instance-based strategies did not make a prediction when a person recalled no or the same number of instances for both sports in a pair comparison. Cue-based strategies did not make a prediction when both sports had identical cue patterns. On average, the social-circle heuristic, the social-circle heuristic_F, and availability-by-recall made predictions in 56%, 57%, and 53% of cases, respectively. In contrast, the take-the-best heuristic, tallying, and Franklin's rule made predictions in 95%, 81%, and 94% of cases, respectively (the recognition heuristic discriminated in 2.8%, $SD = 3.9$, of cases).

One possible reason for the strategies' similar levels of descriptive accuracy is that they often made identical predictions. Table 7 shows that there was, in particular, substantial overlap within the instance-based strategies and within the cue-based strategies. This overlap replicates the results for availability-by-recall and the social-circle heuristic in Study 1. It indicates that predictive information is strongly correlated in natural environments (see Brunswik, 1952), such that strategies using limited search can be on par with compensatory ones (see Diekmann & Rieskamp, 2007; Gigerenzer & Goldstein, 1996). Even between the two strategy classes, the overlap was 60–64%.

Notwithstanding the overlap in predictions, the best instance-based strategy—availability-by-recall—was slightly better at predicting people's inferences than were the best cue-based strategies, Franklin's rule and tallying: $M_s = 72.3\%$ vs. 70.1% , $t(39) = 1.08$, $p = .29$ (Table 6). Finally, the social-circle heuristic (71.9%) achieved nearly the same descriptive accuracy as availability-by-recall (72.3%), but did so with limited search. The heuristic stopped search after the first, second, third, and fourth circles in, on average, 7.3%, 16.2%, 40.3%, and 36.2% of the cases in which it made a prediction (the respective stopping rates for the social-circle heuristic_F were 7.2%, 34.5%, 21.6%, and 36.7%). Thus, the social-circle heuristic was substantially more frugal than availability-by-recall. Averaged across all inferences (in which the strategies did not guess), the social-circle heuristic considered half as many relevant instances as availability-by-recall: $M_s = 1.7$ vs. 3.4 , $t(39) = -6.79$, $p = .001$.

As in Study 1, we used a maximum likelihood approach to classify each participant to one of the six strategies. When no unambiguous classification was possible, the participant was “split up” among

the tied strategies. With one exception, ties occurred only within instance- or within cue-based strategies. Fig. 2 shows the resulting distribution. More than two-thirds of participants were classified as using an instance-based strategy, with the social-circle heuristic (25%) and availability-by-recall (23%) emerging as front-runners. The proportions of participants applying the other strategies were smaller, with 17% each for tallying and Franklin's rule and 8% for take-the-best. As a measure of confidence, we calculated for each strategy classification a Bayes factor quantifying the difference between the best fitting strategy (strategies) and the second-best fitting strategy (strategies).⁸ The larger the difference, the more confident one can be in the classification. A Bayes factor in the range of 1 to 3, 3 to 10, and larger than 10 indicates anecdotal, substantial, and strong evidence, respectively, for the classification (Jeffreys, 1961). Across participants, the median Bayes factor was 10.39, indicating strong evidence. (For a discussion of the possible interplay between instance-based and cue-based inference, see Appendix C.)

In order to examine how important the postulated hierarchy of circles is for the social-circle heuristic's performance, we repeated the classification analysis but reversed the order of circles assumed by the heuristic. Search now proceeded from acquaintances, friends, family to self. In this order, the descriptive accuracy of the social-circle heuristic's dropped considerably, from 26% to merely 13.8% of participants being classified as users of the heuristic.

In sum, in predicting people's inferences about the popularity of sports, instance-based strategies as descriptive models outperformed various cue-based strategies that have been shown to model inferences in a wide range of tasks (Gosling, Ko, Mannarelli, & Morris, 2002; Hammond, 1955; Nosofsky & Bergert, 2007; Pachur & Marinello, 2013; Persson & Rieskamp, 2009; Rieskamp & Hoffrage, 2008). Within the instance-based strategies, about an equal number of participants were best described by the social-circle heuristic and availability-by-recall, respectively.

Tests of response time predictions

The social-circle heuristic assumes a sequential examination of the decision maker's social circles, with search being stopped as soon as the number of instances within a circle discriminates between alternatives. Therefore, the heuristic predicts that the more circles people are required to examine, the more time they need to make an inference. In availability-by-recall, by contrast, tallies are computed across all circles in a compensatory fashion. This strategy therefore does not predict different response times as a function of whether or not a circle discriminates. Similarly, the response times for users of cue-based strategies should not be a function of instances in memory. We tested these response time predictions as follows: Based on the classification

⁸ The Bayes factor (BF) is defined based on the Bayesian Information Criterion (BIC) differences between the best-fitting strategy (strategies) and the second-best fitting strategy (strategies), $BF = \exp(-\frac{1}{2}\Delta BIC)$ (for details, see Wasserman, 2000). The BIC for each strategy is defined as $BIC = C^2 + k \times \log(n)$, with k being the number of free parameters (which equals 1 for all strategies) and n being the number of choices.

Table 6

Percentage of correct predictions (of which of two sports people judge to be more popular) and correct inferences (of which of two sports is more popular) by the six strategies tested in Study 2 for the subset of comparisons in which the respective strategy makes a prediction and the subset in which all strategies make a prediction; standard deviations are in parentheses.

Strategy	Descriptive accuracy		Inferential accuracy	
	The strategy makes a prediction	All strategies make a prediction	The strategy makes a prediction	All strategies make a prediction
Availability-by-recall	71.1% (10.1)	72.3% (10.4)	57.6% (13.6)	59.5% (13.9)
Social-circle heuristic	70.0% (10.0)	71.9% (10.2)	57.1% (13.1)	58.7% (13.7)
Social-circle heuristic _c	69.2% (9.6)	71.5% (10.2)	57.1% (12.9)	58.6% (13.6)
Franklin's rule	66.6% (8.5)	70.1% (12.3)	61.5% (5.0)	66.2% (9.1)
Tallying	69.0% (9.2)	70.1% (12.3)	63.5% (5.5)	66.2% (9.1)
Take-the-best	62.7% (9.1)	65.8% (12.5)	58.5% (7.0)	62.5% (11.3)

described above (Fig. 2), we distinguished between users of the social-circle heuristic, users of availability-by-recall (excluding participants who were equally well described by both strategies), and users of cue-based strategies. To reduce the usual skew in the distribution of response times, they were natural log-transformed (e.g., Whelan, 2008).

We tested the response time predictions using mixed-effects linear models, with 'participants' as random effect and 'circle' as fixed effect. In this analysis, it is important to take into account simple 'symbolic distance effects' (e.g., Moyer & Bayer, 1976): because of the skewed frequency distribution of the sports' popularity (Table 4), comparisons in which an early circle discriminates will likely be those with rather large objective differences: People make these comparison more swiftly than comparisons between objects with smaller differences (e.g., Schweickart & Brown, in press). To control for this possible confound between circle and objective frequency difference, we entered the difference in the actual frequencies between the two sports as a covariate (fixed effect).

Fig. 3 shows the estimated mean response times as a function of the circle at which the social-circle heuristic predicts search to be terminated, separately for the three groups of participants. For participants classified as users of the social-circle heuristic there was a main effect of circle, $F(3, 1240.3) = 3.12, p = .025$, such that response time increased across circles 1 to 4, consistent with the sequential search policy of the heuristic. The earlier the heuristic predicted search to be stopped, the faster the response time. The other strategies predicted no such dependency between circle and response time. Indeed, we obtained only a marginally significant effect of circle for participants classified as users of availability-by-recall, $F(3, 911.3) = 2.25, p = .081$, and no effect for participants classified as users of a cue-based strategy, $F(3, 2813.0) = 1.38, p = .246$.⁹

Fig. 3 also shows that, perhaps surprisingly, users of the social-circle heuristic tended to show slower response times than users of a compensatory instance-based strategy or of cue-based strategies. Interestingly, the same result has also been found among users of lexicographic strategies in cue-based inference (Bröder & Gaissmaier, 2007; Persson & Rieskamp, 2009). One possible explanation is that lexicographic search requires the selection of some information and the inhibition of other information, and that such control processes may incur costs in terms of time (Khader et al., 2011).

How accurate are the strategies' inferences?

We next computed for each participant and based on the information he or she provided—that is, cue values, cue directionality, estimated cue validities, instances and contact frequencies—how accurately instance-based and cue-based strategies predicted the relative popularity of the sports (Table 4), separately for each participant. As in

⁹ The difference in actual frequencies had a significant effect on response time for the users of availability-by-recall, $F(1, 907.8) = 25.7, p = .001$, and the users of cue-based strategies, $F(1, 2818.2) = 25.3, p = .001$, and a marginally significant effect for the users of the social-circle heuristic, $F(1, 1222.5) = 2.76, p = .097$.

Study 1, we focused on cases where the strategies made an unambiguous prediction. The two most accurate strategies were tallying and Franklin's rule, with 64% and 62% correct inferences, respectively (Table 6). The best cue-based strategy, tallying, was significantly more accurate than the best instance-based strategy, availability-by-recall (Kruskal–Wallis test: $z = -2.38, p = .02$). As in Study 1, availability-by-recall (58%) and the social-circle heuristic (57%) reached similar levels of accuracy, despite the latter being more frugal. The same pattern of results was found when the analysis was based on those inferences in which all strategies made a prediction (Table 6).

Summary

The results of Studies 1 and 2 offer some evidence for a noncompensatory and limited instance-based search policy. In Study 2, a quarter of respondents could be classified as users of the social-circle heuristic (with the circles ordered by altruism). The noncompensatory processing of instances among users of the social-circle heuristic was also evident in an analysis of response times. The more circles needed to be examined, the longer it took for respondents to arrive at an inference. Despite foregoing exhaustive search, the social-circle heuristic often reached the same decision as availability-by-recall, which searches social circles exhaustively. This convergence is an important result. It shows that strategies with very different search policies can achieve similar levels of performance in inferring environmental quantities such as social statistics. One possible reason is information redundancy in people's social environments—a property that can be exploited by simple heuristics. This result fits into a growing body of evidence demonstrating that reliance on (relatively) small samples can produce surprisingly competitive judgments and choices (see Fiedler & Kareev, 2006; Gigerenzer & Goldstein, 1996; Hertwig & Pleskac, 2008; Pachur, 2010; but see Juslin, Fiedler, & Chater, 2006).

To our knowledge, our investigation is the first to compare the inferential accuracy of instance-based and cue-based strategies. Why did we find compensatory cue-based strategies to be more accurate than compensatory and noncompensatory instance-based strategies (Table 6)? One possible reason is that the knowledge recruited by the cue-based strategies refers mainly to generic facts that most

Table 7

Proportion of identical predictions of each pair of strategies for the subset of cases for which all strategies made a prediction in Study 2.

Strategy	1	2	3	4	5	6
1. Availability-by-recall	–					
2. Social-circle heuristic	0.97	–				
3. Social-circle heuristic _c	0.94	0.94	–			
4. Franklin's rule	0.64	0.64	0.62	–		
5. Tallying	0.64	0.63	0.62	1	–	
6. Take-the-best	0.62	0.61	0.60	0.86	0.86	–

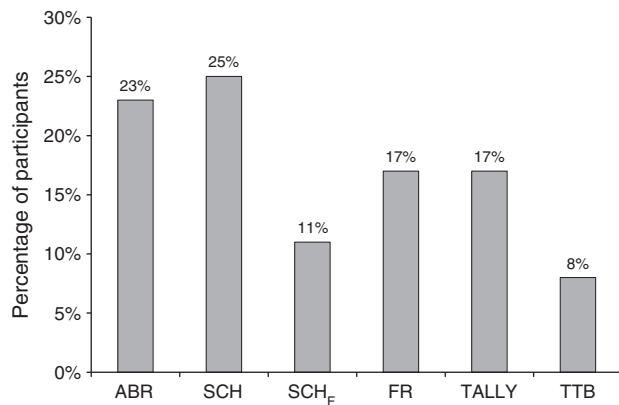


Fig. 2. Individual classification of the 40 participants as following availability-by-recall (ABR), the social-circle heuristic (SCH), the social-circle heuristic_F (SCH_F), Franklin's rule (FR), tallying (TALLY), or take-the-best (TTB) (no participant was classified as guessing).

participants agree upon (e.g., whether a sport is a ball sport), whereas the instance-based strategies recruit idiosyncratic knowledge (whether any of one's friends are members of a basketball club). As a consequence, the predictions of the instance-based strategies are more variable across participants, which can decrease their accuracy.¹⁰ Nevertheless, we caution against overgeneralizing the superiority of the cue-based strategies to other environments. For instance, as Table 3 shows, the circle validities in the cancer mortality environment are substantially higher than those in the sports popularity environment, and the advantage for cue-based inference may therefore be smaller or even disappear there. Furthermore, compensatory cue-based inferences are not consistently more accurate than are noncompensatory cue-based inferences (e.g., Katsikopoulos, Schooler, & Hertwig, 2010).

Study 3: which environmental properties determine the performance of the social-circle heuristic?

In Studies 1 and 2, the social-circle heuristic proved to be able to compete with a compensatory instance-based strategy in inferring environmental frequencies. We now analyze this ability in more detail by studying the ecological rationality (Todd et al., 2012) of the social-circle heuristic. The notion of ecological rationality rests on the assumption that "intelligent behavior in the world comes about by exploiting reliable structures in the world" (Todd et al., 2012, p. viii), thereby directing researchers' focus to the match between mental mechanisms and environmental properties. Rather than merely observing to what extent a particular strategy fails or succeeds across different environments, the goal is to understand which environmental properties a strategy exploits and how they foster or hamper its performance. This approach has led to insights of how simple mental mechanisms, if matched to the right environment, can be surprisingly accurate (see Goldstein & Gigerenzer, 2002; Simon, 1990).

We used computer simulations to examine the impact of two important properties of real-world social environments on the performance of the social-circle heuristic. The first ecological property is the skewness of the frequency distribution across the event categories. Many environmental quantities follow highly skewed distributions (see, e.g., Clauset, Shalizi, & Newman, 2009; Newman, 2005), where very few objects largely dominate all others. Indeed, the two

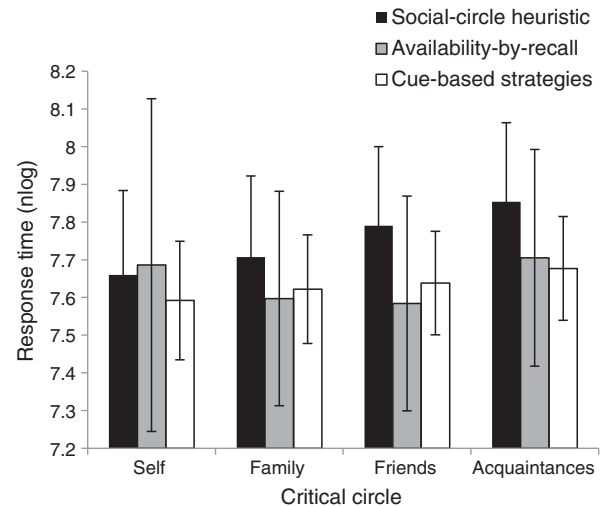


Fig. 3. Estimated mean response times (based on the mixed-effects linear models) as a function of the number of circles examined prior to an inference (according to the social-circle heuristic) in Study 2, separately for participants classified as using the social-circle heuristic, availability-by-recall, and cue-based strategies. Error bars indicate the standard errors of the mean.

social environments of Studies 1 and 2—cancer mortality and popularity of sports—also display very skewed frequency distributions (the latter to a somewhat lesser extent than the former). Hertwig, Hoffrage, and Martignon (1999) and Hertwig, Hoffrage, and Sparr (2013) showed that skewed distributions are conducive to the performance of a noncompensatory cue-based heuristic across social and nonsocial environments, respectively. Does the same hold for a noncompensatory instance-based heuristic?

The second ubiquitous property of social environments is spatial clustering of instances: People tend to know and interact with others who have similar characteristics (see, e.g., McPherson, Smith-Lovin, & Cook, 2001), even causing genetic homogeneity in friendship networks (Fowler, Settle, & Christakis, 2011). Because of contagion or genetic inheritability, many diseases are more prevalent among people who frequently interact with each other. Does such spatial autocorrelation, relative to randomly distributed events, foster or hamper the performance of the social-circle heuristic relative to a compensatory instance-based strategy?

Our computer simulation involved the social-circle heuristic and availability-by-recall. The task of the strategies was to infer which of two event categories occurs more frequently in the population. To our knowledge, this is the first analysis of the ecological rationality of noncompensatory processing of instance knowledge.

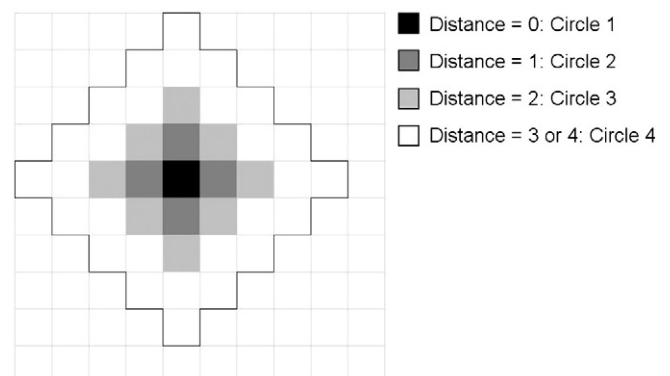


Fig. 4. Representation of the population in the computer simulation in Study 3 (here simplified as a 10 × 10 population).

¹⁰ Our data offer some support for this hypothesis. The average (across participants) standard deviation of the strategies' predictions (averaged across strategies) was lower for the cue-based strategies ($M = .51$) than for the instance-based strategies ($M = 0.79$).

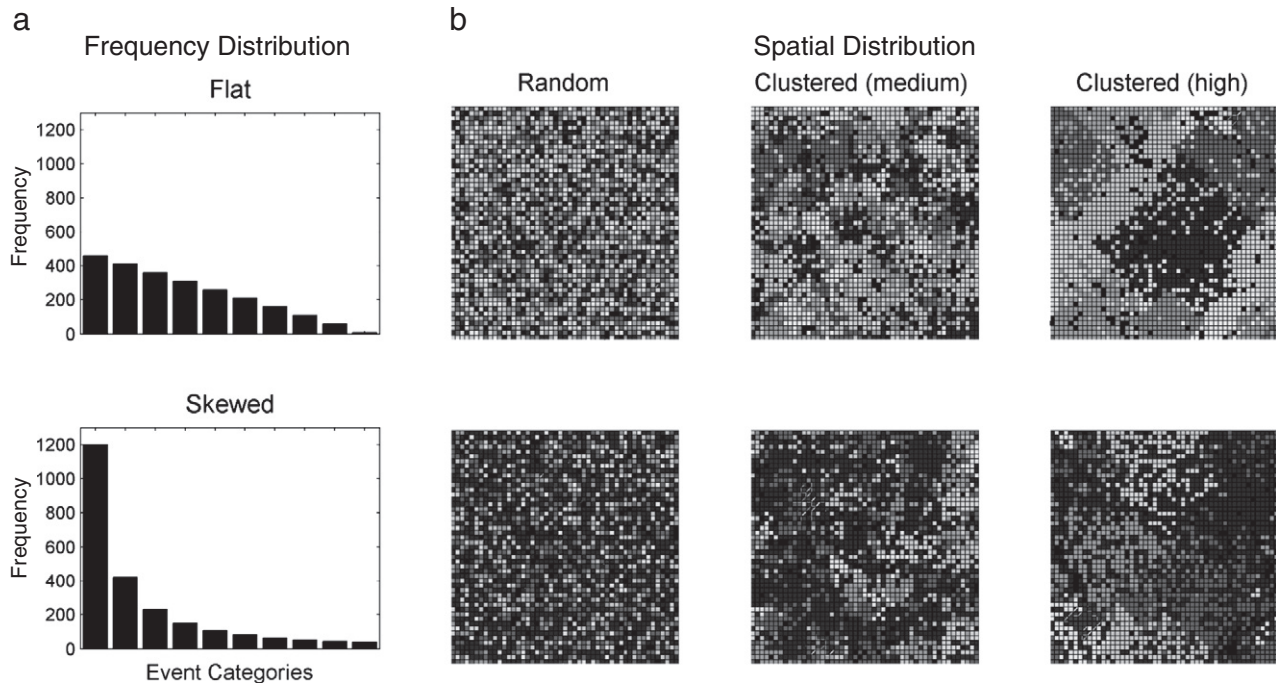


Fig. 5. (a) Frequency distributions of the 10 event categories in the flat and skewed environments investigated in Study 3, and (b) visual illustration of the spatial distributions that resulted from combining flat and skewed environments with different degrees of spatial clustering.

The strategies

We simulated a social environment in which a population of agents was represented as a two-dimensional grid, with each cell representing an agent. To implement the social-circle heuristic, we divided the social network of each agent into four circles. As shown in Fig. 4, the circles were defined in terms of the distance, d , to the focal circle (circle 1), which consisted of the agent itself ($d = 0$). Circle 2 included all four agents with $d = 1$; circle 3, all eight agents with $d = 2$; and circle 4, all 28 agents with $d = 3$ or $d = 4$. An agent's social network thus consisted of 40 other agents, allowing her to sample up to 41 agents (including herself). The size of the circles grew with the distance from the network center (circles 1–4 comprise 1, 4, 8, and 28 agents, respectively), consistent with analyses of actual social networks (Zhou et al., 2005). The social-circle heuristic searched circles 1 through 4 and terminated search when one circle discriminated. Consequently, it probed between a minimum of 1 and a maximum of 41 agents. The compensatory availability-by-recall, in contrast, always probed all 41 agents.

The environments

In each environment we created, there were 10 event categories that differed with regard to their frequency in the population. Occurrences of the event categories were distributed across a population of 2500 agents represented in a 50×50 grid (a common size in agent-based simulations; e.g., Rands, Pettifor, Rowcliffe, & Cowlishaw, 2004). Any given agent was an instance of at most one event category. We examined the strategies' performances in a total of six different environmental conditions, each of which represented a combination of the two properties under consideration: type of frequency distribution (flat vs. skewed) and type of spatial distribution of the instances (random, medium clustering, or high clustering).

Flat versus skewed frequency distributions

The two frequency distributions are shown in Fig. 5a. In the flat environment, the distribution was relatively even and followed a linear function, decreasing smoothly from the most to the least frequent

event category. In the skewed environment, few event categories occurred very frequently and the majority of event categories occurred relatively infrequently (for a related analysis using a real-world frequency distribution, see Pachur, Hertwig, & Rieskamp, 2013). Skewed distributions are often described in terms of a power function (see e.g., Hertwig, Hoffrage, & Sparr, 2013; Newman, 2005). Therefore, we used a power function to generate the distribution in the skewed environment.¹¹

Random versus clustered spatial distribution

Spatial clustering was manipulated by varying the probability p with which the status of an agent (i.e., whether or not it was an instance of a particular event category) affected the status of her neighbors. Specifically, p was a function of the distance, d , between an agent and her neighbor: $p(d) = f(d/5, \mu, \delta)$, where f was a folded cumulative normal distribution with $\mu = 0$. We varied δ to implement three levels of clustering: no clustering ($\delta = 0$), medium clustering ($\delta = 0.7$), and high clustering ($\delta = 7$).

Fig. 5b shows examples of distributions in the six conditions that result from crossing the two types of frequency distributions (flat vs. skewed) with the three levels of clustering. To quantify the degree of clustering, we calculated for each environment the resulting clustering coefficient, expressing the probability that an agent's direct neighbor (defined as having $d = 1$) was an instance of the same event category. The average clustering coefficient in each of the six conditions is shown in Table 8.

The population of agents was represented as a toroidal grid (i.e., a matrix where each cell is connected to the cell on the same row or column, respectively, on the opposite side of the matrix). In each of

¹¹ For both the flat and the skewed environments, the respective functions were adjusted such that the distributions could be well represented in the population of 2500 agents. In the flat environment, the frequency of the x th category (with categories ordered from largest to smallest) was determined by the linear function $f(x) = 10 + x \times 50$. In the skewed environment, the frequency of the x th category was determined by the power function $f(x) = x^{-1.5} \times 1200$. Summing across the resulting frequencies in the event categories, the total number of all agents who were an instance of one of the 10 event categories was 2350 (94% of the population) in the linear environment and 2394 (95.8%) in the skewed environment.

Table 8
The resulting clustering coefficient (CC) in the six environments in Study 3 as well as frugality, proportion of cases in which the strategy had to guess, and accuracy (i.e., proportion of correct inferences) excluding guessing cases for the six environments, separately for the social-circle heuristic and availability-by-recall.

Frequency distribution	Spatial distribution	CC	Social-circle heuristic				Availability-by-recall			
			Frugality		Guessing	Accuracy (no guessing)	Frugality		Guessing	Accuracy (no guessing)
			Relevant instances	Sample size			Relevant instances	Sample size		
Flat	Random	.13	1.78	12.33	0.03	0.76	8.47	41	0.09	0.86
	Clustered (medium)	.36	1.78	18.67	0.16	0.72	9.41	41	0.17	0.74
	Clustered (high)	.58	1.70	25.44	0.44	0.71	14.65	41	0.44	0.71
Skewed	Random	.28	1.62	15.95	0.08	0.79	9.13	41	0.14	0.87
	Clustered (medium)	.47	1.66	21.45	0.26	0.78	11.03	41	0.27	0.80
	Clustered (high)	.57	1.65	25.29	0.46	0.81	15.91	41	0.46	0.82

the six conditions, the simulation was based on 1,000,000 runs. Specifically, each distribution was repeated 10,000 times, and, for each resulting environment, 100 initial agents were sampled randomly. Using the social-circle heuristic or availability-by-recall, each agent inferred in a total of 45 pair comparisons (i.e., all pairwise comparisons of the 10 event categories) which of two event categories is more frequent in the population. If a strategy could not make an inference based on the instances in an agent's network, a random guess was implemented. The strategies' accuracies were defined as the percentage of correct inferences (i.e., the event category inferred to be more frequent was indeed more frequent in the population).

Results

We analyzed the strategies' (a) accuracies in the different conditions and (b) frugality (i.e., number of agents looked up). In addition, we examined the incremental value of further sampling more generally, because a key difference between the social-circle heuristic and availability-by-recall consists in limited versus more extended sampling; specifically, we compared (c) the circle validities and (d) sample size–accuracy trade-off functions between the different environments.

Accuracy

Fig. 6 plots the strategies' accuracies. Three main results emerged. First, for both the compensatory availability-by-recall and the noncompensatory social-circle heuristic, accuracy declined with more clustering in the environment. This is primarily because, with more clustering, both strategies were increasingly unable to discriminate between two event categories and therefore resorted to guessing

(Table 8). Second, the accuracy of both strategies was a function of the combination of skewness and clustering: With no clustering, availability-by-recall performed better in a flat than in a skewed environment (83.1% vs. 81.7%); for the social-circle heuristic, the opposite held (75.2% vs. 76.3%). With a medium level of clustering, however, both strategies performed better in a skewed than in a flat environment (availability-by-recall: 71.9% vs. 69.4%; social-circle heuristic: 70.4% vs. 68.3%). This pattern became even more pronounced with a high level of clustering (availability-by-recall: 66.2% vs. 61.2%; social-circle heuristic: 65.9% vs. 61.1%). Third, the amount of clustering also had a crucial impact on the degree to which the social-circle heuristic was able to compete with availability-by-recall: Without clustering (events distributed randomly), availability-by-recall by far outperformed the social-circle heuristic (averaged across flat and skewed environments, the margin was 6.7 percentage points). With a high level of clustering, this advantage essentially disappeared (the margin shrank to 0.1 and 0.3 percentage points in the flat and the skewed environments, respectively). As Table 8 shows, the impact of clustering on accuracy persisted even when we excluded cases for which an inference was determined randomly (because strategies could not discriminate between event categories).

Frugality

Availability-by-recall searches exhaustively (within the defined social network), whereas the social-circle heuristic limits search. Table 9 reports how often the heuristic stopped search at various circles (i.e., the stopping rates). For instance, in the flat environment with no clustering, in 20% of cases where it did not have to guess, the social-circle heuristic stopped search after considering only the self circle and, as reported in Table 8, made an inference based on a sample of, on average, fewer than two relevant instances (this percentage held across all six conditions). Availability-by-recall, by contrast, made an inference based on a sample of, on average, between 8.5 and 16 relevant instances. Overall, the social-circle heuristic probed (to find relevant instances) an average of between 12 and 25 agents, and thus examined between 29% and 61% of the number of agents probed by availability-by-recall (41 agents). The noncompensatory social-circle heuristic is thus substantially more frugal than the compensatory availability-by-recall strategy, and the differences in frugality depend partly on the degree of clustering in the environment.

Why do skewness and clustering foster the social-circle heuristic's competitiveness?

Using less information than availability-by-recall, the social-circle heuristic approximates the former's accuracy in clustered environments. Why is that? To address this question, we analyzed the circle validities. As Table 9 shows, circle validity is affected by the structure of the environment. In flat environments, circle validity increases across circles; limiting search thus compromises accuracy here. The benefit of more search depends, however, on the degree of clustering:

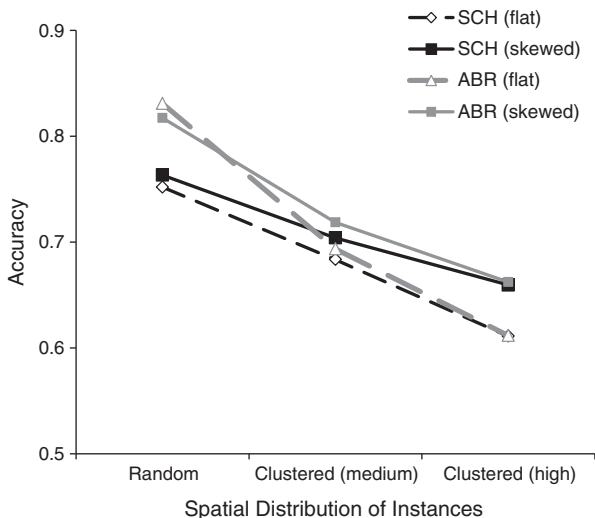


Fig. 6. Accuracy (defined as the proportion of correct inferences) of the social-circle heuristic (SCH) and availability-by-recall (ABR) in the six simulated environments.

Table 9

Average circle validities, discrimination rates (DR), and stopping rates (SR) for each of the four circles in the six environments of the computer simulation for the social-circle heuristic (Study 3); circles 1, 2, 3, and 4 contain 1, 4, 8, and 28 agents, respectively.

Frequency distribution	Spatial distribution	Circle	Validity	DR	SR
Flat	Random	1	0.69	0.19	0.20
		2	0.73	0.48	0.39
		3	0.75	0.65	0.27
		4	0.77	0.87	0.14
	Clustered (medium)	1	0.69	0.19	0.23
		2	0.71	0.41	0.31
		3	0.72	0.57	0.23
		4	0.73	0.80	0.24
	Clustered (high)	1	0.70	0.19	0.33
		2	0.71	0.32	0.26
		3	0.71	0.40	0.18
		4	0.71	0.54	0.23
Skewed	Random	1	0.83	0.19	0.21
		2	0.84	0.44	0.34
		3	0.82	0.58	0.25
		4	0.80	0.82	0.21
	Clustered (medium)	1	0.83	0.19	0.26
		2	0.84	0.38	0.28
		3	0.82	0.50	0.20
		4	0.79	0.71	0.26
	Clustered (high)	1	0.83	0.19	0.35
		2	0.84	0.33	0.29
		3	0.84	0.40	0.16
		4	0.81	0.52	0.20

The more clustered a flat environment is, the lower the increase in validity across circles, and therefore the lower the extent to which more search pays off. In skewed frequency distributions, the link between sample size and circle validity even seems reversed. Seemingly defying the law of large numbers, the validity of circle 1 ($n = 1$) always exceeds that of circle 4 ($n = 28$), independent of clustering.

How can this counterintuitive phenomenon, already observed in Studies 1 and 2 (Table 3), be explained? Skewed frequency environments are skewed because one event category clearly outnumbers others. In such environments, the comparisons in which the smaller circles (e.g., circle 1) discriminate between two event categories are usually “easy” ones—that is, comparisons that involve the most frequent event category and where the difference between the objective frequencies in the population is very large. Such cases are unlikely to be harmed by sampling error. More “difficult” comparisons, that is, those between event categories of medium or low frequencies, usually cannot be determined based on the initial circles—because they are too rare to occur in the “self” or “family” circle—and are therefore delegated to later, larger circles, where the risk of sampling error is lower. In flat environments, by contrast, early (and small) circles will more often also discriminate between rather difficult comparisons, and sampling error will often impair the ability to decide accurately in these comparisons.

The analysis of circle validities also helps to understand why the social-circle heuristic is less vulnerable to the effect of clustering than is availability-by-recall (Fig. 6). For flat environments, the validity of the more peripheral circles, representing larger samples, decreases with a higher level of clustering (e.g., the validity of circle 4 decreases from .77 to .71; Table 9). That is, for samples of size $n > 1$, the statistical principle that event categories that are more frequent in a sample are, *ceteris paribus*, also more frequent in the population is increasingly less valid as the degree of spatial clustering increases. For illustration, consider a sample consisting of three agents, two of which are instances of event category A and one of category B. The occurrence of two instances of A in the sample is less likely to be an indicator that this category is more frequent in the population if instances of an event category occur in clusters than if they do not. If the sample space consists of one agent only (circle 1), by contrast,

the degree of clustering does not affect the extent to which the content of a sample is indicative of the frequency distribution in the population: A single agent is more likely to be an instance of a frequent event category than of a less frequent one, irrespective of whether or not it occurs in a cluster. Consequently, the accuracy of the social-circle heuristic (which often relies on a sample size of $n = 1$) is less affected by clustering than is that of availability-by-recall (which relies on larger sample sizes).

How much information is gained with increasing sample size?

According to the statistical law of large numbers, the accuracy of estimates of population parameters increases with sample size. However, the extent to which “more is better” depends on the structure of the environment. In some environments, a substantial increase in sample size does not necessarily yield a large increase in inferential accuracy. This, in turn, enables the social-circle heuristic to compete with availability-by-recall. To quantify the relationship between the utility of more sampling and ecological structure, we determined sample size–accuracy functions (using computer simulations) and compared them across the six environments. Specifically, we computed inferential accuracy for sample sizes from 1 to 500. For each size, we drew the respective number of agents according to their distance to an initial, randomly selected agent (with close agents sampled first; from agents with the same distance, we sampled randomly the required number). Based on the resulting sample tallies, it was inferred which of two event categories was more frequent in the population (for each of the 45 pair comparisons). When sample tallies did not discriminate, an inference was made randomly. We conducted this analysis across the same six environmental conditions simulated previously. In each condition, 10,000 distributions were generated, and for each distribution 100 initial agents were randomly drawn from the population (yielding a total of 1,000,000 runs per condition).

Figs. 7a and b show the results. First, in all environments, accuracy gain as a function of increasing sample size is subject to a diminishing return (see Hertwig & Pleskac, 2008, 2010). Second, how quickly the return diminishes depends on the environment: As Fig. 7a shows, in all three skewed frequency distributions (gray lines), accuracy increases considerably more steeply at very small sample sizes (and then quickly levels off) than in the flat frequency distributions (black lines). These results further help to understand the contingent competitiveness of the social-circle heuristic: Its policy of limited search is more suitable in environments with skewed frequency distributions.

The social-circle heuristic does not mechanically draw small samples, however. It adjusts sampling conditionally upon the evidence encountered. If an inference can be made based on an early circle, sample size will be small. If an early circle does not render an inference possible, search will be extended. Does this conditional sampling policy—more sampling occurs only if more frugal sampling does not permit an inference—pay off relative to an unconditional sampling policy (i.e., always drawing a fixed sample size)? To find out, we compared the social-circle heuristic’s accuracy with the accuracy achieved on the basis of a fixed (i.e., unconditional) sample size. The black dots in Fig. 7b indicate the heuristic’s performance as a function of the average number of agents it recruited to make an inference in the respective environment (Table 8). The dot lies consistently above the curve for unconditional sampling. This means that with the same average (but otherwise variable) sample size, the heuristic achieves higher accuracy than does an unconditional sampling policy.

In sum, the simulation results suggest an accuracy–effort trade-off in environments with no or medium clustering: The smaller the number of agents that are probed, the less accurate the inferences (Fig. 6). In spatially clustered environments, however, limited search does not inevitably come at the price of a decrease in accuracy: Limited search and accuracy go hand in hand insofar as frugality can guard against drawing erroneous inferences from samples whose representativeness

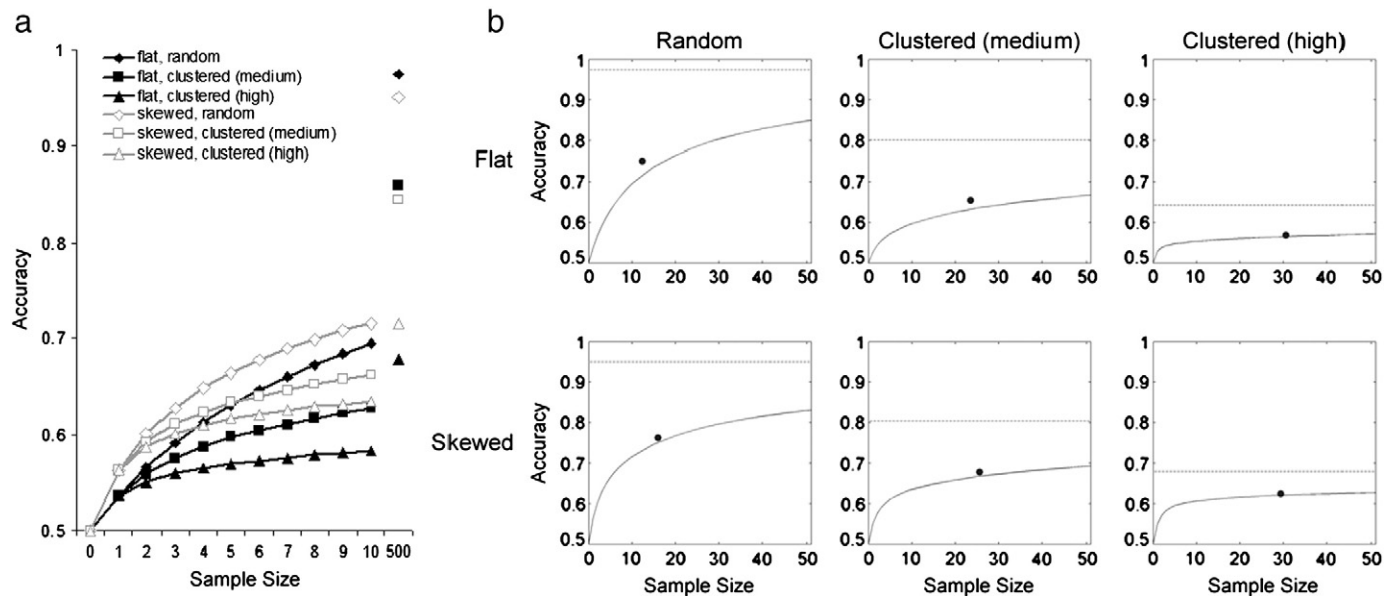


Fig. 7. Accuracy in inferring which of two event categories was more frequent in the population as a function of sample size (i.e., number of agents probed for relevant instances) in the six conditions in the computer simulation (Study 3). The six conditions result from combining two types of frequency distribution (flat vs. skewed) and three degrees of spatial clustering (random vs. medium clustered vs. highly clustered).

is compromised in clustered environments. To analyze a heuristic's ecological rationality is to determine the environmental properties that aid or ail performance. We now can characterize (some) of these properties (Fig. 6). Limited search can compete with exhaustive search in medium and highly clustered environments. In contrast, limited search is far less accurate than exhaustive search when there is no clustering and when the frequency distribution is flat. These regularities also explain why, for the skewed natural environments examined in Studies 1 and 2, the social-circle heuristic proved able to approximate the performance of availability-by-recall.

General discussion

The ability to infer social statistics in the population—such as the frequency of specific behaviors, opinions, and characteristics—is crucial in a world in which others can represent both our fiercest competitors and our closest allies. We investigated to what extent people making such inferences rely on a heuristic that recruits instances and processes them in a noncompensatory fashion. In so doing, we extended the study of noncompensatory heuristics, which have received much attention in cue-based inference (Gigerenzer et al., 1999), to instance-based inference. The social-circle heuristic is the first model of a noncompensatory heuristic that exploits knowledge of instances to infer social statistics (for an illustration of how the heuristic may be applied to the formation of norms and attitudes, see Pachur, Hertwig, et al., 2013). The heuristic assumes that search for instances in memory proceeds sequentially along the hierarchical circle structure of people's social networks (see also Hills & Pachur, 2012), with the circles' boundaries serving as stopping rules.

Our goals were to study the social-circle heuristic from both (a) a descriptive and (b) a prescriptive point of view. Investigating inferences in two different real-world environments, we found that a sizeable proportion of people are well described by the social-circle heuristic. Using response time as a process measure, we further observed that people classified as users of the social-circle heuristic responded more slowly, the more circles the heuristic needed to examine (Fig. 3); we found no such regularity for users of other strategies. Gauging the inferential accuracy of the social-circle heuristic relative to a compensatory strategy, we found that although the heuristic considered, on average, only about half as many instances as

availability-by-recall, it drew the same inference 97% of the time (Study 2). Using computer simulations, we identified two environmental characteristics—clustering and skewness—as key determinants of the competitiveness of the social-circle heuristic. To our knowledge, this is the first analysis of the ecological rationality of an instance-based heuristic. The results also qualify the frequent assumption of a domain-general accuracy–effort trade-off, often believed to be one of the few general laws of the human mind.

Next, we discuss the differential levels of accuracy observed in Studies 1 and 2, the risks and potential benefits of relying on small samples of instances, and the relationship between the social-circle heuristic and memory.

Determinants of accuracy in instance-based inference

Across Studies 1 and 2, the ability of the social-circle heuristic and availability-by-recall to accurately infer social statistics varied widely. In the cancer mortality environment (Study 1), availability-by-recall and the social-circle heuristic yielded 71.4% and 71.9% correct inferences, respectively (when they made an unambiguous prediction); in the popularity of sports environment (Study 2), their accuracy was merely 58% and 57%, respectively. Our ecological analyses (Study 3) offer possible explanations for these variations.

One possible explanation is environmental clustering. According to our simulation results, the performance of both the social-circle heuristic and availability-by-recall strongly declines with increasing degree of spatial clustering (Fig. 6). Therefore, one may speculate that the spatial distribution of cancers is less clustered than that of sports membership. As membership of a sports club almost by definition leads to strong clustering (i.e., one member knows many others who do the same sport), a high level of clustering in this domain is indeed likely. A less likely explanation for the accuracy differences is the skewness of the frequency distribution. Although Fig. 6 shows that skewness can affect accuracy, its impact is much smaller than that of clustering. Moreover, the frequency distributions of cancers and sports are similarly skewed (Fig. 8), with fitted power-law exponents of -1.41 (cancer) and -1.36 (sports), respectively.

Another possible factor behind the different levels of accuracy is the amount of instance knowledge. People retrieved many more instances in Study 2 (sports) than in Study 1 (cancers). Counterintuitively, this

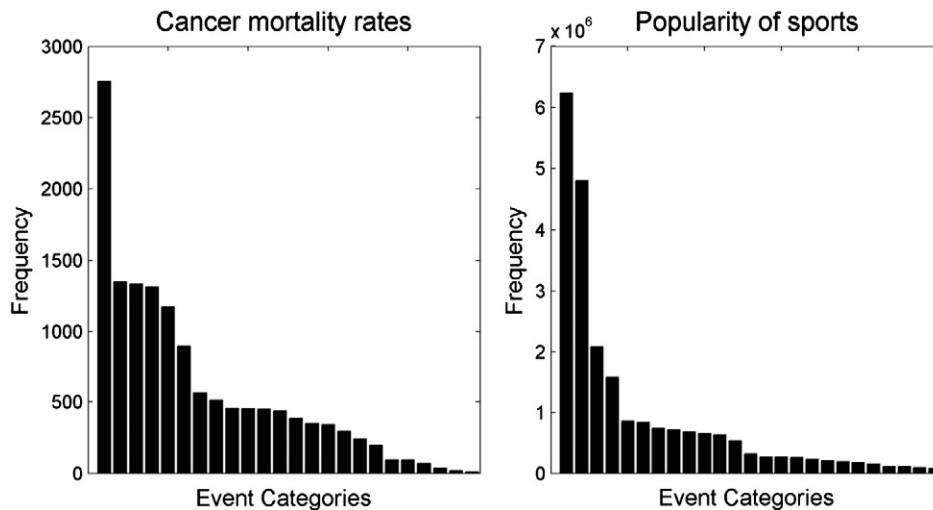


Fig. 8. Frequency distribution of the event categories in the cancer data set and sports data set used in Studies 1 and 2, respectively.

may have reduced the strategies' accuracy in Study 2. The reason is that retrieval of more instances may render the strategies applicable in more pair comparisons—and thus also in pair comparisons that are “difficult.” To test this hypothesis, we calculated, for each participant and separately for the two environments, the median effect size in those pair comparisons for which people's instance knowledge discriminated. Specifically, we determined the effect size measure h for comparing proportions (Cohen, 1992). The average effect size (across participants) was in fact higher for the cancer environment (Study 1) than for the sports environment (Study 2), $M_s = .28$ ($SD = .11$) versus $.17$ ($SD = .06$), $t(20.1) = 4.11$, $p = .001$. That is, the frequency differences (in the population) between the event categories for which participants' instance knowledge discriminated were larger—and inferences were thus “easier”—in the environment in which people knew fewer instances.

Limited search for instances: benefits and boundaries

Recently, it has been argued that reliance on small samples can have a number of important benefits, such as the early detection of useful binary correlations (e.g., Fiedler & Kareev, 2006; Kareev, 2000, 2005; but see Gaissmaier, Schooler, & Rieskamp, 2006; Juslin & Olsson, 2005), the amplification of mean differences and, by extension, the easing of choice difficulty (Hertwig & Pleskac, 2008, 2010). Availability-by-recall relies on samples retrieved from a person's social network. Relatedly, Galesic et al. (2012) proposed a model that assumes that people infer population statistics by relying on a sample of their personal social network. This model can explain seemingly contradictory patterns in social comparison, such as self-enhancement (i.e., evaluating oneself as better than others) and self-depreciation (i.e., evaluating oneself as worse than others). Specifically, it suggests that self-enhancement will occur when the distribution of the general population is J-right shaped (i.e., most people are doing well) and that self-depreciation will occur when the distribution is J-left shaped (i.e., most people are doing badly).

How beneficial or dangerous is availability-by-recall's and the social-circle heuristic's focus on personally experienced instances? It has been argued that “disproportionate exposure, memorability, or imaginability of various events” (Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978, p. 551) can distort estimates of event frequencies. Yet this argument builds on the assumption that the search space in memory extends far beyond a person's social network to include a virtual circle, populated with incidents conveyed through the mass media (Lichtenstein et al., 1978). No doubt, augmenting the search space in memory by a virtual circle comes at the price of systematic error because potential news or entertainment items are selected for

their potential to captivate an audience (see, e.g., Combs & Slovic, 1979). In contrast, sampling only within one's social network—although it constrains the sample size—safeguards people against the media's selection of rare, vivid, dramatic, emotional, and sensational events (for a further discussion, see Pachur et al., 2012; Hertwig et al., 2005).

However, reliance on small samples also exacts risks. Our ecological analyses suggest that one such risk is that of miscalibration due to “clumpiness” in time or space. Illnesses, for instance, often occur in spatial patches or clusters (“hot spots”)—for example, leukemia near nuclear installations, or increased rates of diseases in underserved areas (e.g., Antunes & Waldman, 2002). Similarly, many diseases occur disproportionately in particular age groups. A recent study of the then 229 confirmed human cases of avian influenza type A (H5N1, or “bird flu”) found numerous confirmed cases among children and young adults, with relatively few cases among older adults (Smallman-Raynor & Cliff, 2007). Figs. 6 and 7 show that clustering compromises the accuracy of both compensatory and, though to a less extent, noncompensatory instance-based strategies relying on samples from a person's social network.

Benjamin and Dougan (1997) and Benjamin, Dougan, and Buschena (2001) found that people's estimates of various mortality risks were more in line with event frequencies in their own age cohort than with those in the general population. Interestingly, we observed the same tendency. Respondents' inferences about the popularity of sports (Study 2) were somewhat better tuned to the frequencies in their age cohort (i.e., number of club members aged 27 years and younger) than to the population frequencies. Specifically, participants' accuracies were higher when we used cohort rather than population frequencies as a benchmark, 64.3% versus 62.9%, $t(39) = 1.7$, $p = 0.05$ (one-tailed). But even this sampling bias may be a blessing in disguise. Despite the common notion that the “world is a village,” people typically do not navigate in all social spheres. Therefore, as Benjamin and Dougan argued, accurately estimating social statistics in the population at large may be less important than estimating the social statistics in one's proximate environment.

When do people refrain from instance-based inference?

The results reported here and in Hertwig et al. (2005) suggest that people often rely on instance-based strategies, compensatory and noncompensatory alike. But any heuristic or class of heuristics has boundary conditions. What are those of the social-circle heuristic? First, as with cue-based heuristics (e.g., recognition heuristic, take-the-best), people are unlikely to resort to an inferential strategy if they have direct knowledge about the criterion (see Gigerenzer,

Hoffrage, & Kleinbölting, 1991; Pachur & Hertwig, 2006). To illustrate, when a person happens to know that soccer is the most popular sport in Germany, it follows logically that soccer is more popular than any other sport (irrespective of any instance knowledge). Second, people may rely less on instance knowledge sampled in their social environment if they gauge their proximate social environment to be highly unrepresentative. For instance, when judging the relative frequency of various professions, a carpenter working with other carpenters will hardly infer that carpenters are ubiquitous (e.g., Oppenheimer, 2004; but see Hamill, Wilson, & Nisbett, 1980; Nisbett & Borgida, 1975).

Third, how people generalize from their own status to the population may also depend on the desirability of the characteristic in question. For instance, people tend to see themselves as below-average on highly undesirable characteristics (e.g., watching less trashy TV than the average person) and above-average on highly desirable characteristics (e.g., being less likely to become unemployed than the average person; for an overview, see Chambers, 2008). Cases in which people systematically consider themselves as below average or above average are known as *false uniqueness effects*. In these cases, people seem to see a *negative* link between their own status and that of others. For instance, if a person is not drawn to trashy TV shows, she may infer that many people in the population are. One way to accommodate this phenomenon into the social-circle heuristic would be to assume that, for highly desirable or undesirable characteristics, people reverse the predictive direction of the self circle (i.e., being an instance of an event makes it less likely that it is frequent in the population).

Fluency-based versus instance-based inferences

Availability-by-recall and the social-circle heuristic took their inspiration from the availability heuristic (Tversky & Kahneman, 1973); they represent two ways to turn this heuristic into a computation model. Another strategy related to the availability heuristic is the fluency heuristic (e.g., Hertwig, Herzog, Schooler, & Reimer, 2008; Schooler & Hertwig, 2005). Despite a family resemblance, the fluency heuristic is distinct from availability-by-recall and the social-circle heuristic. It relies on retrieval fluency, defined as the speed with which the name of an event category (or object) is retrieved and recognized. In contrast, availability-by-recall and the social-circle heuristic both entail the retrieval of individual *instances* within the event category. This different input to the strategies has predictable consequences. One is that, unlike the instance-based strategies, the fluency heuristic can distinguish between two event categories (e.g., ovarian cancer vs. bladder cancer) even when no single instance can be retrieved (as long as one category name is retrieved more fluently).

To what extent might people's inferences in Studies 1 and 2 have been driven by the fluency heuristic? As we did not measure people's recognition speed, we are only able to answer this question indirectly. One answer is that past contests between the fluency heuristic and availability-by-recall have consistently suggested that the latter is the superior descriptive model (Hertwig et al., 2005; Pachur et al., 2012; see also Hilbig, Erdfelder, & Pohl, 2011). A second answer concerns the response-time pattern in Fig. 3. For users of the social-circle heuristic, response time increases systematically as a function of the number of circles probed. This pattern is not easily explained by the fluency heuristic.

Another way in which people may make use of mental fluency is by assessing how easy or difficult it is to retrieve instances (e.g., Schwarz et al., 1991). As we did not measure 'ease of retrieval' or manipulate ease and the number of retrieved instances orthogonally, we cannot model their respective impact here. It is possible that people inferring social statistics take into account not just the total number of instances they obtained from their proximate social world, but also the meta-cognitive feeling of ease of retrieval. Yet, because retrieval of instances in the context of the social-circle heuristic is limited to a person's

proximate social world, it is also possible that fluency differences are minor and thus difficult to discern. The link between the social-circle heuristic and fluency is worth modeling in future investigations.

Frequency and/or recency

Two key factors structuring memory and retrieval are frequency and recency of the encoded material (Anderson & Milson, 1989).¹² Interestingly, the distinction between frequency and recency has some parallels with the two social circle definitions we considered. Specifically, the altruism structure—family, friends, and acquaintances—may align with the frequency distribution of contacts accumulated across the lifetime. The frequency-of-contact structure, by contrast, may reflect the frequencies of contacts in a person's more immediate past; it thus could be closely associated with the recency of a person's contacts. Assuming this mapping, we can speculate that the poor performance of the social-circle heuristic_c may indicate that recency is a weaker factor driving the accessibility of instances than is (cumulative) contact frequency. On the other hand, in Hills and Pachur's (2012) investigation of social memory, altruism and frequency of contact predicted retrieval probability about equally well. These inconsistencies suggest that the memory dynamics behind instance-based inferences are not yet fully understood. Future studies could therefore attempt—if possible—to experimentally manipulate frequency and recency in a person's mnemonic repository of instances, and examine their respective impact on the performance of the social-circle heuristic.

Conclusion

Two commandments that are often considered characteristic of rational judgment are exhaustive search ("Thou shalt find all the information available") and compensation ("Thou shalt combine all pieces of information"; Gigerenzer et al., 1999, p. 83). Most psychological models of instance-based inference explicitly or implicitly implement these commandments. Our results suggest that—as in cue-based inference—some people rely instead on limited and ordered search in instance-based inference from memory. Moreover, we identified two environmental properties—spatial clustering and skewness of the frequency distribution—that foster the ability of limited search to compete with exhaustive search. By modeling and analyzing noncompensatory inferences based on instances, our investigations represent a first step toward generalizing the study of ecological rationality (e.g., Todd et al., 2012)—so far limited to the context of cue-based inference—to the noncompensatory processing of knowledge of social instances.

Appendix A. Maximum likelihood classification procedure

The maximum likelihood classification procedure works as follows (e.g., Pachur & Marinello, 2013; for a model recovery analysis, see Pachur & Aebi-Forrer, *in press*). For each of the k candidate strategies and each individual participant, the likelihood of the observed inferences was determined. For this purpose, we determined the G^2 measure (see Burnham & Anderson, 1998), defined as:

$$G^2 = -2 \sum_{i=1}^N \ln [f(y_i|k)], \quad (A1)$$

where $f(y_i|k)$ is the likelihood function that expresses the probability of inference y for item i given strategy k . For those items where strategy k permits an unambiguous prediction, an inference in line with

¹² Pachur, Schooler, and Stevens (2013) showed that these two factors also play a role in the contact dynamics of a person's social environment. Specifically, it was found that the frequency and the recency of previous encounters with a person independently predict the extent to which one is likely to have contact with that person again.

strategy k is predicted with a probability of $(1 - \varepsilon_k)$, where ε_k is a constant application error. Accordingly, an inference inconsistent with strategy k is predicted with probability ε_k . The optimal estimate for the application error ε_k (i.e., the one that leads to the best fit in terms of G^2) can be determined by the proportion of inferences not in line with the strategy's predictions:

$$\varepsilon_k = \frac{n_k}{n}, \quad (\text{A2})$$

where n refers to the number of items for which the strategy makes an unambiguous prediction, and n_k refers to the number of inferences of these n items that are in line with the prediction of strategy k . For all strategies, G^2 is computed based on the items where all strategies make an unambiguous prediction. A participant is classified as using the strategy with the highest likelihood—that is, the lowest G^2 value. If the G^2 of the best-fitting strategy equals or is higher than the G^2 assuming guessing (i.e., $\varepsilon = 0.5$ on all items where all strategies make an unambiguous prediction), then the participant is classified as guessing.

Appendix B. Cue-based strategies tested in Study 2

Take-the-best

The take-the-best heuristic (Gigerenzer & Goldstein, 1996) assumes that a person searches for the most valid cue for judging the relative frequency of two event categories and examines whether the two event categories differ on this cue. The validity v_i of a cue i is defined as $v_i = R_i/D_i$, where R_i is the number of correct inferences by cue i , and D_i is the number of pairs where the value of cue i differs between event categories (in the reference class). The cue validity expresses a cue's conditional probability of making a correct inference given that it discriminates (i.e., that one event category has a positive cue value and the other a negative cue value). If the most valid cue discriminates, a user of take-the-best will terminate search and make an inference accordingly; otherwise, the second most valid cue will be examined, and so on. By using this stopping rule, the heuristic can draw inferences without having to look up the values of all available cues. The heuristic is *noncompensatory*: No amount of evidence contained in the cues not looked up (i.e., cues with a validity below the first cue that discriminates) can reverse the inference based on the discriminating cue.

Tallying

The tallying heuristic represents a very different hypothesis about how people process cues. Tallying searches for all cues in an unspecified order and then infers that the event category with the higher number of positive cue values is the more frequent one. Various versions of tallying with different stopping rules have been discussed in the literature: Either all M cues are looked up, or only the subset of m significant cues is looked up (i.e., search is stopped after m cues, where $1 < m \leq M$; Dawes, 1979). Here we assume that tallying looks up all cues. Unlike take-the-best, tallying is a compensatory heuristic: A negative value on one cue can be compensated for by a positive value on another cue.

Franklin's rule

Like tallying, Franklin's rule looks up the values of all cues. It then weighs the cues by their validity (i.e., the value of cue i is multiplied by v_i), and sums up the weighted values. It infers that the event category with the higher sum is the more frequent one.

Appendix C. Instance-based versus cue-based inference: how do they relate?

In the model comparison in Study 2, we assumed for the sake of conceptual clarity that people used either an instance-based or a cue-based strategy. In reality, this strict assumption may not hold. For instance, a person whose tallies of instances do not discriminate between the event categories may switch to cue-based inference instead of guessing (see Table 1). To address this possibility, we examined how the 27 participants in Study 2 who were classified as users of an instance-based heuristic (availability-by-recall or one of the social-circle heuristics) proceeded when their tallies of instances failed to discriminate between a pair of sports. For those cases, we now classified respondents as "guessers" or "cue users," based on the maximum likelihood method used in Studies 1 and 2, and found that slightly more participants were classified as cue users (15; 56%) than as guessers (12; 44%). This finding is consistent with the idea that some people proceed from an instance-based to a cue-based approach when the instance-based strategy fails to enable an inference (see Fig. 1).

References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036–1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.
- Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703–719.
- Antunes, J. L. F., & Waldman, E. A. (2002). Trends and spatial distribution of deaths of children aged 12–60 months in São Paulo, Brazil, 1980–98. *Bulletin of the World Health Organization*, 80, 391–398.
- Benjamin, D. K., & Dougan, W. R. (1997). Individuals' estimates of the risks of death: Part I—A reassessment of the previous evidence. *Journal of Risk and Uncertainty*, 15, 115–133.
- Benjamin, D. K., Dougan, W. R., & Buschena, D. (2001). Individuals' estimates of the risks of death: Part II—New evidence. *Journal of Risk and Uncertainty*, 22, 35–57.
- Bond, C. F., Jr., & Brockett, D. R. (1987). A social context personality index theory of memory for acquaintances. *Journal of Personality and Social Psychology*, 52, 1110–1121.
- Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science*, 21, 471–475.
- Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multi-attribute decisions. *Psychonomic Bulletin and Review*, 14, 895–900.
- Bröder, A., & Schiffer, S. (2003). "Take-the-best" versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General*, 132, 277–293.
- Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making*, 19, 361–380.
- Brunswik, E. (1952). *The conceptual framework of psychology*. *International encyclopedia of unified science, vol. I, no. 10*, Chicago, IL: University of Chicago Press.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193–217.
- Bundesamt für Statistik (2004). *Todesursachenstatistik: Ursachen der Sterblichkeit 1999 und 2000 [Statistics on causes of death: Causes of death in 1999 and 2000]*. Neuchâtel, Switzerland: Bundesamt für Statistik.
- Burnham, K. P., & Anderson, D. R. (1998). *Model selection and inference: A practical information-theoretic approach*. New York, NY: Springer.
- Chambers, J. R. (2008). Explaining false uniqueness: Why we're both better and worse than others. *Social and Personality Psychology Compass*, 2, 1–17.
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51, 661–703.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Combs, B., & Slovic, P. (1979). Causes of death: Biased newspaper coverage and biased judgments. *Journalism Quarterly*, 56(837–843), 849.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 571–582.
- De Schryver, M., Vandist, K., & Rosseel, Y. (2009). How many exemplars are used? Explorations with the Rex Leopold I model. *Psychonomic Bulletin and Review*, 16, 337–343.
- Dhimi, M., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, 130, 959–988.
- Dieckmann, A., & Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Memory and Cognition*, 35, 1801–1813.
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106, 180–209.
- ECDC Rapid Risk Assessment (May 27). *Outbreak of Shiga toxin-producing E. coli (STEC) in Germany*. Stockholm: European Centre for Disease Prevention and Control.
- Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, 73, 221–230.

- Fiedler, K., & Juslin, P. (Eds.). (2005). *Information sampling and adaptive cognition*. New York, NY: Cambridge University Press.
- Fiedler, K., & Kareev, Y. (2006). Does decision quality (always) increase with the size of information samples? Some vicissitudes in applying the law of large numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 883–903.
- Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrinic, C., Kastenmüller, A., Frey, D., et al. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137, 517–537.
- Fiske, A. P. (1995). Social schemata for remembering people: Relationships and person attributes in free recall of acquaintances. *Journal of Quantitative Anthropology*, 5, 305–324.
- Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. H., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Decision Processes*, 43, 75–117.
- Fowler, J., Settle, J., & Christakis, N. (2011). Correlated genotypes in friendship networks. *Proceedings of the National Academy of Sciences*, 108, 1993–1997.
- Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 966–982.
- Galesic, M., Olsson, H., & Rieskamp, J. (2012). Social sampling explains apparent biases in judgments of social environments. *Psychological Science*, 23, 1515–1523.
- Geary, D. C., & Bjorklund, D. F. (2000). Evolutionary developmental psychology. *Child Development*, 71, 57–65.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.). (2011). *Heuristics: The foundations of adaptive behavior*. New York: Oxford University Press.
- Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas. *Psychological Review*, 115, 230–239.
- Gigerenzer, G., Hoffrage, U., & Kleinböting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.
- Goldstein, N. J., Cialdini, R. B., & Criskevicius (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35, 472–482.
- Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109, 75–90.
- Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Judgments of personality based on offices and bedrooms. *Journal of Personality and Social Psychology*, 82, 379–398.
- Hamill, R., Wilson, T. D., & Nisbett, R. E. (1980). Insensitivity to sample bias: Generalizing from atypical cases. *Journal of Personality and Social Psychology*, 39, 578–589.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour, I and II. *Journal of Theoretical Biology*, 7, 1–52.
- Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review*, 62, 255–262.
- Henrich, N., & Henrich, J. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. New York, NY: Oxford University Press.
- Hertwig, R., & Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Social Cognition*, 27, 661–698.
- Hertwig, R., Herzog, S. M., Schooler, L. J., & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1191–1206.
- Hertwig, R., Hoffrage, U., & Martignon, L. (1999). Quick estimation: Letting the environment do the work. In G. Gigerenzer, P. M. Todd, & the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 209–234). New York, NY: Oxford University Press.
- Hertwig, R., Hoffrage, U., & Sparr, R. (2013). The QuickEst heuristic: How to benefit from an imbalanced world. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world*. Oxford, UK: Oxford University Press.
- Hertwig, R., Hoffrage, U., & the ABC Research Group (2013). *Simple heuristics in a social world*. New York, NY: Oxford University Press.
- Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 621–642.
- Hertwig, R., & Pleskac, T. J. (2008). The game of life: How small samples render choices simpler. In N. Chater, & M. Oaksford (Eds.), *The probabilistic mind: Prospects for rational models of cognition* (pp. 209–235). Oxford, UK: Oxford University Press.
- Hertwig, R., & Pleskac, T. J. (2010). How small samples amplify perceived differences in decisions from experience. *Cognition*, 115, 225–237.
- Hilbig, B. E., Erdfelder, E., & Pohl, R. F. (2011). Fluent, fast, and frugal? A formal model evaluation of the interplay between memory, fluency, and comparative judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 827–839.
- Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature*, 14, 53–72.
- Hills, T. T., & Pachur, T. (2012). Dynamic search and working memory in social recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 218–228.
- Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces. *Psychological Science*, 19, 676–682.
- Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Clarendon Press.
- Juslin, P., Fiedler, K., & Chater, N. (2006). Less is more in contingency assessment. Or is it? In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition*. Cambridge, UK: Cambridge University Press.
- Juslin, P., & Olsson, H. (2005). Capacity limitations and the detection of correlations: Comment on Kareev (2000). *Psychological Review*, 112, 256–267.
- Juslin, P., & Persson, M. (2002). PROBABILITIES from EXemplars (PROBEX): A “lazy” algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, 26, 563–607.
- Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. *Life Span Development and Behavior*, 3, 253–286.
- Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, 107, 397–402.
- Kareev, Y. (2005). And yet the small-sample effect does hold: Reply to Juslin and Olsson (2005) and Anderson, Doherty, Berg, and Friedrich (2005). *Psychological Review*, 112, 280–285.
- Katsikopoulos, K., Pachur, T., Machery, E., & Wallin, A. (2008). From Meehl (1954) to fast and frugal heuristics (and back): New insights into how to bridge the clinical-actuarial divide. *Theory and Psychology*, 18, 443–464.
- Katsikopoulos, K. V., Schooler, L. J., & Hertwig, R. (2010). The robust beauty of ordinary information. *Psychological Review*, 117, 1259–1266.
- Khader, P. H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). Memory-based decision making with heuristics involves increased activation of decision-relevant memory representations. *Journal of Cognitive Neuroscience*, 23, 3540–3554.
- Krueger, J. I., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, 67, 596–610.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 551–578.
- Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin*, 102, 72–90.
- Martignon, L., & Hoffrage, U. (2002). Fast, frugal and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52, 29–71.
- McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Milardo, R. M. (1992). Comparative methods for delineating social networks. *Journal of Social and Personal Relationships*, 9, 447–461.
- Moreno, J. L. (1936). Organization of the social atom. *Sociometric Review*, 1, 11–16.
- Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 8, 228–246.
- Newman, M. E. J. (2005). Power laws, Pareto distributions, and Zip’s law. *Contemporary Physics*, 46, 323–351.
- Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 32, 932–943.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115, 39–57.
- Nosofsky, R. M., & Bergert, F. B. (2007). Limitations of exemplar models of multi-attribute probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 999–1019.
- Oppenheimer, D. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science*, 15, 100–105.
- Pachur, T. (2010). Recognition-based inference: When is less more in the real world? *Psychonomic Bulletin and Review*, 17, 589–598.
- Pachur, T., & Aebi-Forrer, E. (2013). Selection of decision strategies after conscious and unconscious thought. *Journal of Behavioral Decision Making*. <http://dx.doi.org/10.1002/bdm.1780> (in press, Advance online publication).
- Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 983–1002.
- Pachur, T., Hertwig, R., & Rieskamp, J. (2013). The mind as an intuitive pollster: Frugal search in social spaces. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 261–291). New York, NY: Oxford University Press.
- Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied*, 18, 314–330.
- Pachur, T., & Marinello, G. (2013). Expert intuitions: How to model the decision strategies of airport customs officers? *Acta Psychologica*, 144, 97–103.
- Pachur, T., & Olsson, H. (2012). Type of learning task impacts performance and strategy selection in decision making. *Cognitive Psychology*, 65, 207–240.
- Pachur, T., Schooler, L. J., & Stevens, J. R. (2013). When will we meet again? Regularities of social connectivity and their reflections in memory and decision making. In R. Hertwig, U. Hoffrage, & the ABC Research Group (Eds.), *Simple heuristics in a social world* (pp. 199–224). New York: Oxford University Press.
- Pachur, T., Todd, P. M., Gigerenzer, G., Schooler, L. J., & Goldstein, D. G. (2011). The recognition heuristic: A review of theory and tests. *Frontiers in Cognitive Science*, 2, 147.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision-maker*. Cambridge, UK: Cambridge University Press.
- Persson, M., & Rieskamp, J. (2009). Inferences from memory: Strategy- and exemplar-based judgment models compared. *Acta Psychologica*, 130, 25–37.
- Pirollo, P., & Card, S. K. (1999). Information foraging. *Psychological Review*, 106, 643–675.
- Rands, S. A., Pettifor, R. A., Rowcliffe, J. M., & Cowlshaw, G. (2004). State-dependent foraging rules for social animals in selfish herds. *Proceedings of the Royal Society B: Biological Sciences*, 271, 2613–2620.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, 127, 258–276.

- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13, 279–301.
- Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311, 854–856.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–627.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61, 195–202.
- Schweickart, O., & Brown, N. R. (2013). Magnitude comparison extended: How lack of knowledge informs comparative judgments under uncertainty. *Journal of Experimental Psychology: General*. <http://dx.doi.org/10.1037/a0031451> (in press, Advance online publication).
- Sedlmeier, P., Hertwig, R., & Gigerenzer, G. (1998). Are judgments of the positional frequencies of letters systematically biased due to availability? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 754–770.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41, 1–19.
- Singer, P. (1981). *The expanding circle: Ethics and sociobiology*. Oxford, UK: Oxford University Press.
- Smallman-Raynor, M., & Cliff, A. D. (2007). Avian influenza A (H5N1) age distribution in humans. *Emerging Infectious Diseases*, 13, 510.
- Statistisches Bundesamt (2002). *Statistisches Jahrbuch 2001 für die Bundesrepublik Deutschland [Statistical yearbook 2001 for the Federal Republic of Germany]*. Wiesbaden, Germany: Author.
- Todd, P. M., Gigerenzer, G., & the ABC Research Group (Eds.). (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281–299.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, 44, 92–107.
- Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record*, 58, 475–482.
- Wilson, T. D., & Brekke, N. C. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116, 117–142.
- Zhou, W.-X., Sornette, D., Hill, R. A., & Dunbar, R. I. M. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B*, 272, 439–444.