On the Psychology of the Recognition Heuristic: Retrieval Primacy as a Key Determinant of Its Use

Thorsten Pachur
Max Planck Institute for Human Development

Ralph Hertwig
University of Basel

The recognition heuristic is a prime example of a boundedly rational mind tool that rests on an evolved capacity, recognition, and exploits environmental structures. When originally proposed, it was conjectured that no other probabilistic cue reverses the recognition-based inference (D. G. Goldstein & G. Gigerenzer, 2002). More recent studies challenged this view and gave rise to the argument that recognition enters inferences just like any other probabilistic cue. By linking research on the heuristic with research on recognition memory, the authors argue that the retrieval of recognition information is not tantamount to the retrieval of other probabilistic cues. Specifically, the retrieval of subjective recognition precedes that of an objective probabilistic cue and occurs at little to no cognitive cost. This retrieval primacy gives rise to 2 predictions, both of which have been empirically supported: Inferences in line with the recognition heuristic (a) are made faster than inferences inconsistent with it and (b) are more prevalent under time pressure. Suspension of the heuristic, in contrast, requires additional time, and direct knowledge of the criterion variable, if available, can trigger such suspension.

Keywords: decision making, ecological rationality, fast and frugal heuristics, memory, recognition

Research on human cognition gives rise to a strange paradox. On the one hand, it has been demonstrated that higher order processes—which allow a person to perform the mental feats that are held to distinguish mankind from all other creatures (e.g., Dawkins, 1989; Dennett, 1996; Mead, 1935)—are subject to a myriad of bounds (e.g., Cowan, 2001; Kahneman, Slovic, & Tversky, 1982). On the other hand, elementary processes such as those involved in perception, memory, and motor coordination—which humans are likely to share with other animals—seem to be accomplished by a fantastically complex machinery. In other words, whereas evolution appears to have equipped humans with prodigious processing capacities for seemingly *elementary* processes, it turned strangely stingy when it reached the crowning faculty distinguishing "us" from "them": our ability for higher order cognitive processes.

Why would evolution allocate resources and processing capacity so asymmetrically—barely limited capacities and powerful processors for elementary processes and a "limited-capacity information processor" (Payne, Bettman, & Johnson, 1993, p. 9) for the high art of reasoning? There are several possible reasons, some more obvious than others. One is the considerable cost involved in growing and maintaining a large, high-energy-expending brain (Martin, 1983). Another is the counterintuitive adaptive benefit of cognitive bounds (e.g., Hertwig & Todd, 2003; Kareev, 2000; but see Juslin & Olsson, 2005; Schooler & Hertwig, 2005). Still another reason is that the higher cognitive processes can escape capacity limits by co-opting more elementary but complex abilities. According to this argument, the human mind represents an adaptive toolbox of simple cognitive strategies that are capable of co-opting automatic and complex evolved (or learned) abilities, thus allowing the conscious machinery to stay lean (Gigerenzer, Todd, & the ABC Research Group, 1999). But co-optation is not the only reason why simple strategies in the mental toolbox can abstain from complexity and yet strive for accuracy. Another reason is that they are assumed to exploit informational regularities in the environment. Good performance thus is also a function of the appropriate mapping of simple inference tools to environments. Ecologically rational behavior arises from selecting the right tool for the right environment, rather than using a single universal inference tool suitable for every situation (for a similar view, see also Payne et al., 1993).

The topic of this article is a key example of an ecologically rational strategy co-opting a complex capacity. The simple *recognition heuristic* (Goldstein & Gigerenzer, 2002) hinges on the vast, sensitive, and reliable capacity for recognition. Arguably the most frugal within the program of fast and frugal heuristics (Gigerenzer et al., 1999), the recognition heuristic makes an inference from systematic patterns of existing and missing knowledge. In this article, we further elaborate on how recognition is used for probabilistic inference. Specifically, we examine recognition's exceptional status due to its mnemonic properties, the heuristic's boundary conditions, and the way recognition links up with other knowledge. We begin by describing the recognition heuristic, the capacity it exploits, and the controversial thesis that recognition gives rise to noncompensatory inferences.

Correspondence concerning this article should be addressed to Thorsten Pachur, Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Lentzeallee 94, Berlin 14195, Germany. E-mail: pachur@mpib-berlin.mpg.de

The Recognition Heuristic: A Tool Co-Opting an Evolved Capacity

You are a contestant on the ABC show *Who Wants to Be a Millionaire*. As your final $1 million question, Regis Philbin asks you: "Which of the following two musicians has as of today sold more albums in the U.S.A.: George Strait or Billy Joel?" What is your answer? If you are American, the question may strike you as quite problematic. You may, for instance, remember that pop legend Billy Joel has won numerous Grammy Awards, was inducted into the Rock and Roll Hall of Fame, and has released several Top 10 albums. At the same time, you may also think of the many platinum albums that country music legend George Strait has earned, not to mention his many American Music Awards and Academy of Country Music honors. If the choice were tough for an American who happens to know all these facts, how difficult would it be for a European, say, a Swiss, who in all likelihood has never heard of George Strait (93% of students at the University of Basel did not recognize his name; Herzog, 2005), let alone his many achievements?

Yet, could it be that the clueless Swiss contestant would be, paradoxically, more likely to hit on the right answer than the clued-up American counterpart? More generally, is it possible that people who know less about a subject nevertheless make more correct inferences than their better-informed counterparts? Indeed, it is possible. If the less-informed person—for instance, the Swiss facing the Billy Joel versus George Strait question—exploited his or her ignorance by using the recognition heuristic, he or she would answer the question correctly (for further details on such a less-is-more effect, see Goldstein & Gigerenzer, 2002). For a two-alternative choice task, such as choosing between Billy Joel and George Strait, the recognition heuristic can be stated as follows:

> If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value with respect to the criterion.

Accordingly, the Swiss contestant would infer that the recognized artist has sold more albums. Having heard of both artists, the savvy American contender, ironically, knows too much to be able to take advantage of the recognition heuristic.

As with other strategies from the mental toolbox, the recognition heuristic can afford to be a simple, one-reason decision-making strategy (Gigerenzer et al., 1999) because it feeds on the outcome of an evolved (and automatized) capacity. In this case, it is the capacity for recognition that enables processes such as face, voice, and name recognition. By co-opting this capacity—that in itself is likely to be a complex ability (e.g., Wallis & Bülthoff, 1999)—the recognition heuristic taxes the cognitive resources only modestly. In addition, the recognition heuristic exploits a frequent informational regularity in the environment: Whether we recognize something is often not random but systematic. Therefore, the recognition heuristic promises to be useful, as Goldstein and Gigerenzer (2002) pointed out, whenever there is a strong correlation—in either direction—between recognition and the criterion (for simplicity, we assume henceforth that the correlation is positive). As the mind-as-an-adaptive-toolbox metaphor implies, people should resort to using other tools if this correlation is weak or even nonexistent (see also Gigerenzer & Goldstein, 1996, p. 653). Before we turn to what we know about how people appear to use recognition knowledge, we first describe how Goldstein and Gigerenzer envisioned its use.

The Noncompensatory Status of Recognition Information: Mixed Evidence

The capacity for recognition is often assumed to have played a pivotal role in a number of adaptive problems, ranging from avoidance of strangers (Scarr & Salapatek, 1970) to avoidance of poisonous food. In these evolutionarily important domains, recognition is typically observed to be used in a noncompensatory way (e.g., Galef, McQuoid, & Whiskin, 1990). In light of its evolutionary history, Goldstein and Gigerenzer (2002, p. 77) referred to recognition as a "primordial psychological mechanism" and proposed that the capacity for recognition is being co-opted for drawing probabilistic inferences in the here and now. The recognition heuristic embodies one mind tool through which this co-optation occurs. Moreover, the same authors assumed that the typically noncompensatory status of recognition information observed in evolutionarily important domains generalizes to probabilistic inferences: "The recognition heuristic is a noncompensatory strategy: If one object is recognized and the other is not, then the inference is determined" (Goldstein & Gigerenzer, 2002, p. 82).

The term *noncompensatory* means that for a decision task that is solved using probabilistic information—cues or attributes—a choice for an object based on one attribute "cannot be reversed by other attributes of the object," that is, the attributes are not integrated into a single judgment (Elrod, Johnson, & White, 2005, p. 2; see also Payne et al., 1993, p. 29). Relatedly, the recognition heuristic is noncompensatory in that it does not allow room for the integration of recognition knowledge with other probabilistic cues: It "relies only on subjective recognition and not on objective cues" (Goldstein & Gigerenzer, 2002, p. 82). This does not mean, however, that no other knowledge—such as direct knowledge of the object's criterion value—can override the verdict of the recognition heuristic. In our view, this very point and, more generally, the meaning of the term noncompensatory have led to some confusion. We return to this shortly.

Since Goldstein and Gigerenzer (2002) proposed the recognition heuristic, numerous studies have demonstrated that recognition or lack thereof is an important piece of information across various inferential tasks.[1] At the same time the assumption that it is used in a noncompensatory way has been vigorously challenged (e.g., Bröder & Eichler, 2006; Newell & Fernandez, in press; Newell & Shanks, 2004; Oppenheimer, 2003; Pohl, 2006; Richter & Späth, 2006). Goldstein and Gigerenzer (2002) originally tested this assumption by pitting recognition information against other, conflicting probabilistic cues. Specifically, American students were tested on their ability to infer which was the larger of two German cities. Goldstein and Gigerenzer found that despite the presence of conflicting useful cue knowledge that participants had learned during

---

[1] Recognition information has been shown to be used across a range of inferential tasks such as the prediction of outcomes at sports events (Pachur & Biele, in press; Serwe & Frings, in press), political elections (Marewski, Gaissmaier, Dieckmann, Schooler, & Gigerenzer, 2005), and the judgment of demographic, geographic, and biological quantities (Pohl, 2006; Reimer & Katsikopoulos, 2004; Richter & Späth, 2006).

the experiment (e.g., that a particular recognized city has no soccer team), on average 92% of inferences were consistent with the recognition heuristic, suggesting that recognition knowledge overrode knowledge of objective probabilistic cues (but see Newell & Fernandez, in press, Experiment 1; Richter & Späth, 2006, Experiment 3).

In an inventive set of studies, Newell and Shanks (2004) extended the test of the recognition heuristic to a situation in which participants learned to "recognize" fictional company names (consisting of nonwords—i.e., none of the names were recognized before the experiment), which were presented repeatedly to them (Bröder & Eichler, 2006, used a similar methodology). Moreover, the validity of the induced recognition was manipulated. In a subsequent judgment task the participants were to infer which of two companies—one recognized, one unrecognized—had the more profitable stock. To aid their decision, people could purchase additional cues in the form of experts' advice. The validity of the cues (i.e., recognition and the recommendations of three advisors) was learned through feedback over the course of the experiment. Consistent with the recognition heuristic, in the majority of choices the recognized company was chosen to be more profitable (88%; see Newell & Shanks, 2004, Table 2). In addition, recognition was frequently (68% of all cases) the only cue used (i.e., no further information was purchased). However, this was only so when recognition was the most valid cue. When it was the cue with the lowest validity, most participants (64%) purchased additional information and, based on the experts' advice, a substantial proportion picked the stock they did not recognize (in 38% of cases). Newell and Shanks (2004) concluded: "We found little evidence suggesting that recognition is treated any differently from other cues in the environment" (p. 932). In their view, recognition is usually integrated with other available cue knowledge (see also Richter & Späth, 2006). On the basis of the observation that knowledge of additional probabilistic cues—participants learned them during the experiment—affected the use of (induced) recognition, Bröder and Eichler (2006) arrived at the same conclusion.

Thus, Newell and Shanks's (2004) findings appear to suggest that people stray from the use of recognition as described in the model of the recognition heuristic. How representative, however, is the context in which their participants found themselves in these studies? They knew that recognition was inferior to all other accessible cues. They knew the context in which they learned to recognize an object. Outside the laboratory one is rarely so clairvoyant. For instance, one is typically not able to pin down and discern between the various contexts in which one may have previously encountered the names of cities, Goldstein and Gigerenzer's (2002) domain of inference. Thus, Newell and Shanks's (and Bröder & Eichler's, 2006) results may in fact demonstrate that an induced sense of recognition—which can be unmistakably traced to one source, the experiment—may not give rise to the same use of recognition as would a naturally evolved sense of recognition. The latter typically cannot be traced exclusively to one specific source. Such an interpretation of their results also conforms with evidence indicating that in making inferences people appear to rely less on subjective assessments of memory (e.g., processing fluency) when they can attribute this memory to the experiment than when such an

explicit attribution is impossible (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989; Oppenheimer, 2004; Schwarz et al., 1991).

The relation of recognition and other knowledge was also the subject of Oppenheimer's (2003) investigation. Unlike Newell and Shanks's (2004) studies, his involved recognition that partly evolved outside the laboratory. Specifically, he presented Stanford University students with pairs of well-known and fictitious cities. Their task was to choose the larger one. The well-known cities were carefully selected such that participants either knew that the city they recognized was relatively small (e.g., Sausalito in his Experiment 1) or they knew that their ability to recognize a city was due to factors other than its size (e.g., Chernobyl in his Experiment 2). In both contexts, Oppenheimer found that recognition information was overruled. The unrecognized fictitious cities were systematically inferred to be *larger* than the recognized cities (i.e., > 50% of the time).

Suspending the recognition heuristic when one explicitly knows that a city is very small, however, does not conflict with the model of the heuristic. In answering questions such as which of two cities is larger, it is plausible to assume that the mind attempts a direct solution by retrieving definitive knowledge about the criterion that gives rise to a *local mental model* (LMM; Gigerenzer, Hoffrage, & Kleinbölting, 1991). In general, an LMM can be successfully constructed if (a) precise figures can be retrieved from memory for both alternatives (e.g., cities), (b) nonoverlapping intervals of possible criterion values can be retrieved, or (c) elementary logical operations can compensate for missing knowledge (e.g., if one city is the largest or the smallest in the set, then any other will by definition be smaller or larger, respectively). An LMM represents a local and direct solution. No use of the probabilistic cue–environment structure is made, and merely the presented alternatives and their criterion values are taken into account.[2] According to Gigerenzer et al., only if no LMM can be constructed will inductive inferences involving probabilistic cues need to compensate for missing direct knowledge. The recognition heuristic is meant to be one model for such an inductive inference, in which "the criterion is not immediately accessible to the organism" (Goldstein & Gigerenzer, 2002, p. 78; see also Gigerenzer & Goldstein, 1996). Returning to Oppenheimer's (2003) results, one interpretation is that his students did not use the recognition heuristic because they succeeded in constructing an LMM, for instance, by assuming that Sausalito is so small that one can safely deduce that the other city, even if not recognized, is larger. Pohl's (2006) results can be interpreted similarly. Across four studies, he found that the choice of a recognized object depends, sometimes to a great extent, on whether this choice proves to be correct or incorrect. This contingency would arise if direct (valid) criterion knowledge were available. Finally, Richter and Späth's (2006, Experiment 1)

---

[2] Because people's knowledge is imperfect, it is not guaranteed that LMMs yield accurate solutions. Moreover, factors such as forgetting and fluctuations in retrieval performance can result in intervals of criterion values rather than precise point estimates.

findings are also consistent with the view that criterion knowledge mediates the use of the recognition heuristic.[3]

In our view, Newell and Shanks's (2004) and Oppenheimer's (2003) studies identified two important situations in which people clearly do not use the recognition heuristic. First, the heuristic appears not to be triggered or is overruled when recognition knowledge did not evolve naturally or when recognition can be traced to one source that is dissociated from the criterion variable. Second, the heuristic, as an inductive device, will only be used if a direct solution fails. Under these conditions, the evidence suggests that people's inferences are not determined by recognition but by information beyond recognition. But these boundary conditions, in our view, do not warrant the conclusion that recognition information is treated on a par with any other probabilistic information. We submit the thesis that recognition information—independent of its precise confluence with direct and probabilistic knowledge—is not just like "any other" probabilistic cue (Newell & Shanks, 2004, p. 928). Because of its mnemonic properties, recognition has an exceptional status. To appreciate this thesis, let us turn next to research on recognition memory.

Recognition Information: First on the Mental Stage

For more than 30 years, memory researchers have attempted to identify the processes underlying recognition judgments (see Yonelinas, 2002, for an overview). Although there is ongoing debate as to whether recognition is based on a single, global matching process (see Clark & Gronlund, 1996, for a review) or can better be described in terms of a dual-process account (e.g., Jacoby, 1991), there is consensus that two different kinds of information contribute to recognition.[4] One is a global sense of familiarity, which is "generally thought to reflect an assessment of the global similarity between studied and tested items." The other is recollection, entailing "the retrieval of specific information about studied items, such as physical attributes . . . , associative/contextual information . . . , or other source-specifying information" (both quotes are from Curran, 2000, p. 923).

To illustrate, for the task of discriminating between a studied word and a dissimilar nonstudied word, one can often rely primarily on a global sense of familiarity. This global information, however, will not suffice to reject a word, for example, *house*, that was received aurally when the word *house* was initially studied as a written item. To reject the heard word *house*, one has to recollect associative knowledge, namely the modality in which the word was studied. Similarly, in a memory game, it does not suffice to recognize that a currently turned-over card is the counterpart of a card turned over previously. One also has to recollect the position of the previously revealed card.

A key difference between familiarity and recollection is that familiarity enters the mental stage earlier than information accrued by recollection (Gronlund & Ratcliff, 1989; Hintzman & Curran, 1994; McElree, Dolan, & Jacoby, 1999; Ratcliff & McKoon, 1989). This retrieval advantage is taken to indicate that familiarity represents an automatic form of memory, whereas recollection involves an intentional, slow, and effortful retrieval process (Atkinson & Juola, 1974; Jacoby, 1991; Mandler, 1980).

Using the distinction between familiarity and recollection, Payne, Richardson, and Howes (2000) examined the role of rec-

ognition in problem solving. Similarly, we ask whether the same distinction, and in particular, the temporal dissociation between familiarity and recollection, can also be relevant for the recognition heuristic. We believe so. First, recall that Goldstein and Gigerenzer (2002) used the term *recognition* to refer to the discrimination "between the truly novel and the previously experienced" (p. 77). To render this discrimination possible, familiarity information often suffices, and no associative information (episodic or other knowledge associated with the objects) needs to be recollected (as, for instance, in many lexical decision tasks).

Second, and more important, we suggest that the dissociation between familiarity and recollection observed in recognition tasks extends to recognition and other cue knowledge in inference tasks. Specifically, information about an object—including probabilistic cues such as whether a given German city has a soccer team or whether Billy Joel has won numerous Grammy Awards—requires effortful retrieval, just as does recollection of knowledge about the modality in which an item was studied. Recognition knowledge, by contrast and in analogy to familiarity, is provided automatically. As a consequence, recognition is first on the mental stage and ready to enter inferential processes when other probabilistic cues still await retrieval. Henceforth, we refer to these properties as the *retrieval primacy* of recognition information.

Predictions

The notion that recognition has a retrieval primacy has testable implications. In what follows, we elaborate these implications in terms of three predictions.

> *Prediction 1.* Shorter response times are needed for recognition-based inferences. Inferences that agree with the recognition heuristic require less response time than choices that are inconsistent with the recognition heuristic.

This prediction is derived as follows: Information about a global sense of familiarity, which suffices to make a recognition judgment, is available almost immediately. Therefore inferences based on recognition will be made expeditiously. In contrast, inferences inconsistent with the recognition heuristic need to rely on information beyond recognition (unless they are produced by mere guessing), such as associative information (e.g., source information), probabilistic cues, or knowledge of the criterion variable. As the latter typically require effort and time for retrieval, such inferences will, on average, require more time than inferences consistent with the recognition heuristic.

Prediction 1 has an interesting corollary: The longer it takes to arrive at a response, the more likely the response will disagree with the recognition heuristic (provided that the additionally retrieved

---

[3] In the decision task of Richter and Späth's (2006) Experiment 1, participants judged which of two animal species has a larger population. Additional knowledge was assessed by asking participants to indicate whether a species is an endangered one. As endangered species have by definition a small population size, this knowledge represents criterion knowledge.

[4] See Gronlund and Ratcliff (1989) and Clark and Gronlund (1996) for accounts of the possible contribution of these two kinds of information to recognition judgments in (modified) global matching models.

information contradicts the choice determined by recognition). In other words, with increasing response time there will be a monotonic drop in the proportion of inferences consistent with the recognition heuristic. This regularity follows from the fact that the more time elapses, the more knowledge beyond recognition (if available) can be retrieved. Consequently, the longer the response time, the weaker the impact of recognition on the final judgment.

Prediction 2. Time pressure fosters recognition-based inferences. Limited time to make inferences will lead to greater use of the recognition heuristic, and consequently to more inferences consistent with the heuristic.

Prediction 2 is derived as follows: Recognition is assumed to precede the retrieval of other knowledge such as probabilistic cues. Because recognition is available when other knowledge could not yet be accessed, it will have more impact on the inferences when this process is subject to time pressure.

The final set of predictions concerns the notion of the adaptive use of the recognition heuristic. Recognition information is generated automatically and thus cannot be suppressed. Arguably, this retrieval primacy, being a result of our cognitive architecture, holds irrespective of how well recognition tracks the criterion. This raises the question of how people take account of recognition when it poorly predicts the criterion. Will they then resort to other knowledge and other strategies? We conjecture that the notion of the adaptive toolbox strongly implies that people can and will resort to other strategies if reliance on recognition is anticipated to be futile. To render precise predictions possible, we exploit another key property of the recognition heuristic. The recognition heuristic is domain specific; that is, its use will only be successful if recognition is correlated with the criterion. The heuristic's attainable accuracy (i.e., the percentage of correct inferences) in an environment is indexed by the *recognition validity* α, which can be calculated as $\alpha = R / (R + W)$, where R and W equal the number of correct and incorrect inferences, respectively (across all inferences in which one object is recognized and the other is not). Typically, the recognition validity α is calculated for each participant, and the average of these "personal" αs (across participants) is then taken as an indicator of the recognition validity in a domain (see Goldstein & Gigerenzer, 2002, p. 87).

How is the heuristic used in environments in which recognition is but a poor predictor of the criterion? Gigerenzer and Goldstein (1996) acknowledged this situation and argued that "in cases where recognition does not predict the target, [the inference is performed] without the recognition principle" (p. 663). Goldstein and Gigerenzer (2002), however, did not specify any specific threshold of α for the use of the heuristic. Instead, they suggested that as long as α surpasses .5, the heuristic is used, even if conflicting and markedly more valid cues could be retrieved (given the purported noncompensatory use of recognition). In our view, the notion of a match between environments and heuristics implies that people should resort to other knowledge if in a given domain recognition knowledge proves a poor predictor. But how would such an adaptive use be achieved? In particular, given that recognition is likely to precede the retrieval of other knowledge and is impossible to hold back, how can one adjust the heuristic's use? In what follows, we propose three hypotheses (Predictions 3a–3c) for a restrained use of the recognition heuristic in environments with low α.

Prediction 3a: Threshold hypothesis. Users of the heuristic rely invariably on recognition as long as α exceeds a threshold. If a given environment's α is below this threshold, users will not employ the heuristic.

Such a threshold hypothesis is consistent with the observation in previous studies that the mean rates of adherence to the recognition heuristic were consistently high (i.e., around 90%) in spite of highly variable αs (see, e.g., Pachur & Biele, in press; Pohl, 2006; Reimer & Katsikopoulos, 2004; Serwe & Frings, in press).

The threshold hypothesis suggests three testable regularities. First, although (given the currently limited knowledge) we cannot precisely pin down the numerical value of such a threshold, it should be located between .5 and the lowest α observed to date in association with a high adherence rate, namely .7 (Pachur & Biele, in press). Second, the hypothesis predicts two distinguishable clusters of adherence rates: one encompassing high adherence rates (users whose α exceeds the threshold) and another including low adherence rates (users whose α is below the threshold). Third, there should be a strong positive correlation between individuals' αs and their adherence rates (under the assumption that people have some ability to correctly assess the validity of their recognition knowledge for a given inference task).

Prediction 3b: Matching hypothesis. Users of the heuristic follow it with a probability that matches their individual α, the recognition validity.

This hypothesis is inspired by the frequent observation of people choosing the more likely of two events with a probability matching that event's probability of success. Specifically, when people have to choose between two options A and B, and A leads to a success with probability p and B leads to a success with probability $q = 1 - p$, people respond as if they were probability matching. That is, if $p > q$, rather than always choosing A (i.e., probability maximization), they distribute their responses such that A is chosen with a probability of p and B is chosen with a probability of $1 - p$ (e.g., Gallistel, 1990; Vulkan, 2000).

In the context of the recognition heuristic, this hypothesis predicts that people match their use of the heuristic to their individual α. Consequently, the recognized object is chosen to be the larger one with a probability of $p = \alpha$, whereas the unrecognized object is chosen to be the larger one with a probability of $q = 1 - \alpha$. From this, it follows that the proportion of inferences consistent with the recognition heuristic should equal α. That is, like the threshold hypothesis, the matching hypothesis implies a strong correlation between adherence rates and individuals' αs (again assuming some correspondence between people's perceived and actual recognition validity). Pohl (2006) obtained some evidence for such a correlation. In contrast to the threshold hypothesis, the matching hypothesis does not imply two clearly distinguishable clusters of adherence rates but a graded variation in adherence rates as a function of people's α.

Prediction 3c: Suspension hypothesis. The nonuse of the recognition heuristic does not hinge on recognition validity α but on object-specific knowledge that is at odds with recognition.

Such object-specific contradictory knowledge can come in different forms, including (a) source knowledge (i.e., if a person realizes that her recognition of an object is clearly due to a factor other than the object's criterion value, for instance, the presentation of an object within an experiment) and (b) direct conflicting knowledge of an object's criterion value (see Oppenheimer, 2003), which, as pointed out above, allows for the construction of an LMM. Here we focus on the latter. If, for instance, a Stanford University student is asked to judge which city has more inhabitants, Sausalito or Gelsenkirchen, the student might zero in on Gelsenkirchen even though he or she does not recognize this German city. The reason is that the student knows that Sausalito, with around 7,500 residents, is a very small city and therefore he or she suspends the recognition heuristic for this specific inference.

Unlike the first two hypotheses, the suspension hypothesis implies marked variability in the use of the recognition heuristic across objects and across participants. This is because some objects are more likely to be associated with direct knowledge than others (e.g., students of Stanford University are likely to know that Sausalito is comparatively small and may put the recognition heuristic aside in all pairs that involve Sausalito), and some people have direct knowledge where others lack it. By probing for the individual availability of LMMs, we will be able to investigate whether they are likely to co-occur with the nonuse of the heuristic. If they do, there will likely not be a strong link between people's αs and their recognition heuristic adherence, a link obligatory for the other two hypotheses.

Predictions 3a–3c represent three different hypotheses of how users of the recognition heuristic restrain the use of the recognition heuristic. Such restrained use may be particularly apt when α is low. Before we turn to an empirical test of Predictions 1–3, let us address an important objection. Does the notion of the recognition heuristic's adaptive use render it too flexible and, perhaps, unfalsifiable? For two reasons, we do not think so. First, the adaptive use of the heuristic implies that environments with low α are less likely to give rise to its use than environments with medium or high α. Adaptive use is thus tantamount to a robust and systematic pattern of both use and nonuse. Similarly, the specific hypotheses underlying the nonuse of the recognition heuristic define testable constraints. Nonuse is not thought to be random but to manifest itself in predictable and testable ways. By testing such constraints, we follow Newell's (2005) call to further elucidate the boundary conditions of the adaptive toolbox.

The Environment

Two studies tested Predictions 1, 2, and 3. Both studies used variants of the same experimental procedure. Participants were given pairs of infectious diseases, and their task was to choose the more prevalent in each pair. We chose this domain primarily because it requires the retrieval of knowledge acquired outside the laboratory, thus liberating us from using experimentally induced recognition or artificially created environments. Equally important, Prediction 3 requires the study of an environment in which recognition is of comparatively low validity. Conveniently, such an environment is also appropriate to test Predictions 1 and 2. Both necessitate an environment in which at least some of the knowledge people have conflicts with the recognition heuristic. This is likely to happen in an environment with low recognition validity.
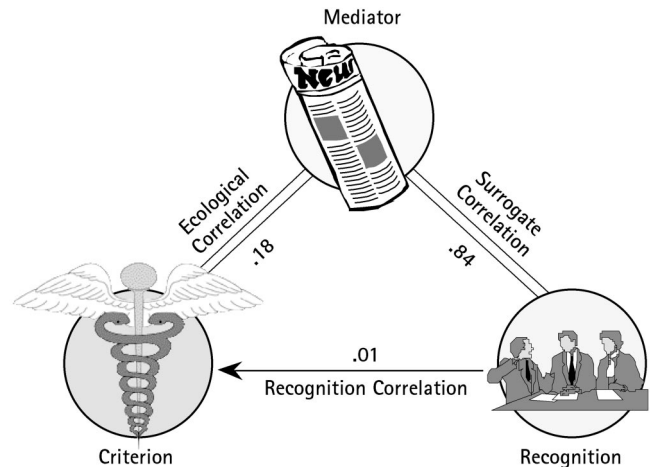

Figure 1. Ecological analysis of recognition. Recognition is highly correlated with media coverage, whereas both media coverage and recognition are uncorrelated with incidence rates of the infectious diseases. Recognition is a very poor indicator of incidence rates of the diseases, because media coverage, acting as mediator, does not (only) reflect the incidence rates (but possibly also the severity of the disease).

The domain of infectious diseases represents such an environment (see Hertwig, Pachur, & Kurzenhäuser, 2005).[5]

Figure 1 depicts the relationships between annual incidence rates of 24 notifiable infectious diseases in Germany, the frequency with which the names of the diseases were mentioned in the media, and *collective recognition* (i.e., the proportion of participants recognizing each infection in Study 1; see Goldstein & Gigerenzer, 2002).[6] The frequencies of mentions in the media, assumed to operate as the mediator between the criterion and recognition, were determined using COSMAS (Corpus Search, Management and Analysis System) I, an extensive data archive of German daily and weekly newspaper articles.[7] We determined the number of times the names of the 24 infections were mentioned and rank correlated these numbers with collective recognition. As Figure 1 shows, media coverage was highly correlated with collective recognition (*surrogate correlation*: $r_s = .84$, $p = .001$), in line with the assumption that recognition is determined by how often the names of infections occur in the environment (for which mention frequency in the media is assumed to be a proxy; Goldstein & Gigerenzer, 2002). In contrast, the correlation between the

---

[5] Hertwig et al. (2005) did not investigate the recognition heuristic directly. However, they found only a modest correlation ($r_s = .23$) between the incidence rates of the diseases and the frequency with which the infections were mentioned in the media, the latter being a strong predictor of recognition (according to Goldstein & Gigerenzer, 2002).

[6] Classified as particularly dangerous, occurrences of these diseases have to be registered. To determine the correct answers, we used statistics prepared by the Federal Statistical Office of Germany and the Robert Koch Institute (e.g., Robert Koch Institute, 2001). To reduce year-to-year fluctuations, we averaged the data across 4 consecutive years (1997–2000).

[7] COSMAS is the largest online archive of German literature (e.g., encyclopedias, books, and newspaper articles; http://corpora.ids-mannheim.de/~cosmas/). Our analysis was based on a total of 1,211 million words.

Table 1
The 24 Infectious Diseases Used as Target Events in Studies 1 and 2

Target event	Study 1 (N = 40)				Study 2 (N = 60)				
	Annual incidence rate	Recognized by % of participants	Proportion of choices in line with RH (M)	n	Recognized by % of participants	Proportion of choices in line with RH (M)	n	Estimated incidence (Mdn)	% of participants with direct knowledge
Poliomyelitis	0.25	100.0	.57	40	100.0	.70	59	50	30.0
Diphtheria	1	97.5	.66	39	98.3	.70	58	500	18.3
Trachoma	1.75	7.5	.76	3	13.3	.49	7	50	5.0
Tularemia	2	2.5	1.00	1	3.3	.57	1	50	1.7
Cholera	3	100.0	.30	40	100.0	.47	59	5	31.7
Leprosy	5	100.0	.15	40	100.0	.37	59	5	30.0
Tetanus	9	100.0	.66	40	100.0	.69	59	500	23.3
Hemorrhagic fever	10	20.0	.76	8	33.3	.82	19	500	6.7
Botulism	15	22.5	.63	8	18.3	.70	10	50	8.3
Trichinosis	22	20.0	.60	9	23.3	.67	13	50	5.0
Brucellosis	23	12.5	.66	5	15.0	.83	8	50	5.0
Leptospirosis	39	7.5	.42	3	25.0	.68	14	50	5.0
Gas gangrene	98	27.5	.38	11	28.3	.65	16	50	11.7
Ornithosis	119	7.5	.54	3	10.0	.79	5	50	5.0
Typhoid and paratyphoid	152	87.5	.46	35	90.0	.77	53	50	16.7
Q fever	179	12.5	.37	5	16.7	.56	9	50	5.0
Malaria	936	100.0	.63	40	100.0	.59	59	500	26.7
Syphilis	1,514	95.0	.59	38	100.0	.76	59	500	21.7
Shigellosis	1,627	5.0	.90	2	20.0	.64	11	50	5.0
Gonorrhea	2,926	95.0	.74	38	96.7	.72	57	5,000	18.3
Meningitis and encephalitis	4,019	97.5	.79	39	91.7	.88	54	5,000	20.0
Tuberculosis	12,619	100.0	.67	40	98.3	.69	58	500	26.7
Viral hepatitis	14,889	90.0	.91	36	86.7	.84	51	5,000	18.3
Gastroenteritis	203,864	85.0	.97	34	96.7	.92	57	200,000	26.7

Note. RH = recognition heuristic.

criterion and the mediator, the *ecological correlation*, was weak ($r_s = .18$, $p = .39$). It is notable that the correlation between collective recognition and the infections' incidence rates turned out to be nil ($r_s = .01$, $p = .95$). That is, the proportion of participants who recognized the infections did not reflect the actual incidence rates of the diseases. Undeniably, recognition is a poor predictor of the criterion in this environment *hostile* to the recognition heuristic.

Study 1: Does the Recognition Heuristic Give Way to Faster Choices?

If one object is recognized and the other is not, the recognition heuristic can determine the choice without searching and retrieving other probabilistic cues about the recognized object. The reversal of a choice determined by recognition, in contrast, requires retrieval of further information (unless the reversal reflects mere guessing). Hinging on this difference, Prediction 1 states that inferences agreeing with the recognition heuristic require less response time than choices that are inconsistent with the recognition heuristic. Study 1 tests this prediction as well as the predictions following from the three candidate hypotheses of how a restrained use of the recognition heuristic in environments with a low α is implemented (Prediction 3).

Method

Participants and design. Forty students from Free University (Berlin, Germany) participated in the study (27 women and 13 men, mean age =

24.2 years), which was conducted at the Max Planck Institute for Human Development in Berlin. They were presented with pairs of names of infectious diseases and asked to choose the infection with the higher annual incidence rate in a typical year in Germany (henceforth *choice task*). They also indicated which of the infections they recognized (henceforth *recognition task*). All were paid for participating. Half of the participants received a flat fee of 9 Euros ($11.80 U.S.) and monetary incentive in the form of a performance-contingent payment. Specifically, they earned 4 cents (5 cents U.S.) for each correct choice and lost 4 cents for each wrong one. The other half of participants received a flat fee of 10 Euros ($13.10 U.S.). Participants were randomly assigned to one of the four conditions of a 2 (recognition test before/after the choice task) × 2 (monetary incentive/no incentive) design, with 10 participants in each condition.

Materials. For the choice task, we used all 24 infectious diseases (see Table 1) and generated all 276 possible pairs, which were presented in 12 blocks (each containing 23 pairs). Both the order in which the 276 pairs of infections appeared and the order of the infections within each pair were determined at random for each participant. The recognition task comprised all 24 infections.

Procedure. After reading an introductory text explaining the relevance of accurate judgments of the frequency of dangerous infectious diseases, participants read the following instructions:

> We ask you to judge the annual frequency of occurrence of different types of infections in Germany. . . . Each item consists of two different types of infections. The question you are to answer is: For which of the two infections is the number of new incidents per year larger?

Pairs of the names of infections were displayed on a computer screen. Participants were asked to indicate their choice by pressing one of two keys. In addition, they were instructed to keep the index fingers of the right

and left hands positioned on the keys representing the right and left elements in the pair of infections, respectively, for the entire duration of one block. They were encouraged to respond as quickly and accurately as possible (although they were not told that their response times were being recorded). The time that elapsed between the presentation of the infections and participants' keystrokes was measured. Each choice began with the presentation of a fixation point (a cross in the center of the screen), followed after 1,000 ms by the infections. The names appeared simultaneously (left and right from the fixation point) and remained on the screen until a response was given. Participants were informed that once the response key was pressed, their choice could not be reversed. After each response, the screen remained blank for 1,000 ms. To accustom participants to the procedure, we asked them to respond to 10 practice trials. The practice trials consisted of 10 pairs randomly drawn from the 276 pairs of infections, which were used again in the main task.

After conclusion of the choice task, half of the participants took the recognition task. In this task, the 24 infections were presented in alphabetical order on a questionnaire, and participants indicated whether they had heard of the infection before the experiment. Half of participants took the recognition test prior to the choice task. On average, the complete session lasted around 60 min.

Results

First, we describe the obtained inferences in more detail. On average, participants scored 60.9% (SD = 5.6%) correct. Neither incentives, $F(2, 35)$ = 0.43, p = .66, nor the order of the recognition task, $F(2, 35)$ = 1.24, p = .30, had a significant effect on the level of accuracy or the proportion of choices in line with the recognition heuristic. Therefore, we pooled the data for the following analyses. On average, participants recognized 58% (range = 37.5%–95.8%) of the 24 infections. Recognition rates are listed in Table 1. The frequency of recognized infections did not increase significantly when the recognition task succeeded the choice task, $t(38)$ = 1.31, p = .20. Across all participants and items, the recognition heuristic was applicable in almost half of all pairs (M = 48.5%, SD = 8.0%). Finally, the average recognition validity α was .60 (SD = .07); thus recognition knowledge (measured in terms of α) proved, on average, modestly helpful in inferring disease incidence rates. The average knowledge validity β—expressing the accuracy in cases when both diseases were recognized—was .66 (SD = .08).

Did the recognition heuristic predict people's inferences? For each participant, we computed the percentage of inferences that were in line with the recognition heuristic among all cases in which it could be applied (i.e., where one infection was recognized and the other not). The mean percentage of inferences in line with the recognition heuristic was 62.1% (Mdn = 62.7%). The present adherence rate is markedly lower than in Goldstein and Gigerenzer (2002), who found proportions of 90% and higher (in a task involving choosing the larger of two cities). Hence, we succeeded in investigating an environment in which people did not obey the recognition heuristic in a substantial portion of their judgments, thus creating a test bed for Prediction 1.

Were inferences in accordance with the recognition heuristic made faster (Prediction 1)? We analyzed the response times by taking choices rather than participants as the unit of analysis. Figure 2 shows the 25th, 50th, and 75th percentiles of the response-time distribution, separately for inferences consistent and inconsistent with the recognition heuristic. In line with Prediction 1, we found that response times for inferences that agreed with the
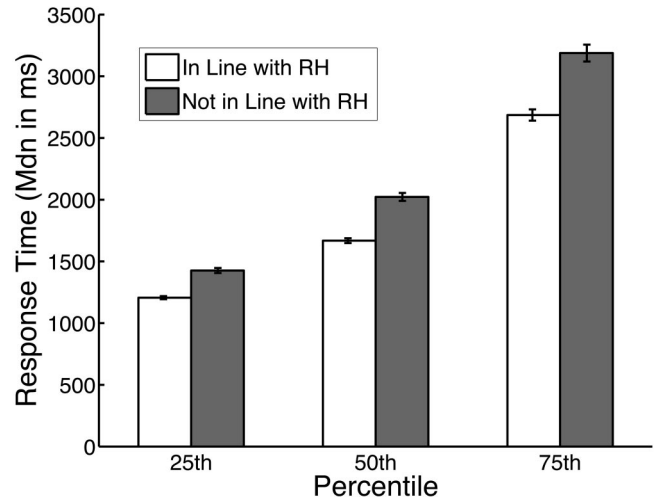

Figure 2. Distribution of the response times of choices where the recognition heuristic (RH) was applicable. The 25th, 50th, and 75th percentiles of response times are shown as a function of whether or not the choice was in line with the recognition heuristic. The error bars indicate standard errors. Because standard errors are not defined for percentiles, we used the standard deviations of the sampling distribution of the 25th, 50th, and 75th percentiles (Howell, 2002). These standard deviations were obtained using a bootstrapping procedure based on 10,000 draws with replacement.

heuristic were substantially shorter at each of the three percentiles than choices conflicting with the heuristic. For instance, the response times for the 50th percentile were 1,668 ms and 2,022.5 ms, respectively. Inferences in line with the recognition heuristic also took less time than inferences in which the recognition heuristic was not applicable, with medians of 2,032 ms and 1,953.5 ms when both diseases were unrecognized and recognized, respectively.

Prediction 1 was also confirmed by a second analysis, in which response times were natural log-transformed to reduce the skewness of the data. Figure 3 compares the average response times for inferences consistent and inconsistent with the recognition heuristic. Inferences that conflicted with the recognition heuristic took longer (M = 7.7, SD = 0.6) than those consistent with the recognition heuristic (M = 7.5, SD = 0.6), $t(5353)$ = 10.8, p = .001, Cohen's d = 0.30. As Figure 3 also shows, the response times for incorrect inferences were markedly longer than for correct inferences, irrespective of whether they agreed with the recognition heuristic. This pattern reflects a typical finding in the memory literature, especially in tasks in which the overall accuracy is low (e.g., Ratcliff & Smith, 2004).

Thus, in support of Prediction 1, inferences that agreed with the recognition heuristic were made faster than those that went against it. This observation supports the notion that recognition information outruns other inferential information. The decision not to use the recognition heuristic appears to exact the cost of longer response times.

Which hypothesis captures the restricted use of the recognition heuristic best (Predictions 3a–3c)? As observed earlier, the recognition heuristic accordance is markedly lower in the infectious diseases environment than in other environments previously studied. At the same time, we have obtained support that recognition is

the first cue on the mental stage, so people somehow managed to escape from relying too much on this instantaneous information. Thus, we now have an opportunity to investigate which of the proposed hypotheses—the threshold, the matching, or the suspension hypothesis—best captures people's restrained use of the recognition heuristic in this environment. We begin with the threshold hypothesis, according to which the average recognition heuristic accordance represents the combination of two clusters of adherence rates: first, the high rates of participants who invariably rely on the heuristic because their individual recognition validity α exceeds the critical threshold, and second, the low rates of those who never use the heuristic because their α is below threshold. Figure 4 plots each participant's adherence rate (i.e., the percentage of inferences that agreed with the recognition heuristic among all cases in which it could be applied) as a function of that participant's recognition validity α. Each point in Figure 4 represents 1 participant. As can be seen, the distribution of adherence rates does not resemble that implied by the threshold hypothesis. Rather than showing two clusters of adherence rates—one cluster of high rates and one of low rates—the actual rates varied continuously between 35.8% and 95.1%.

Looking at the data in Figure 4 also renders possible a test of the matching hypothesis. According to this hypothesis, the user of the recognition heuristic uses it with a probability corresponding to his or her recognition validity α. On an aggregate level, the proportion of choices following the recognition heuristic indeed closely matched the average α: .62 versus .60. As Figure 4 shows, however, when individual adherence rates and αs are considered, this match proves spurious. Rather than being lined up along the diagonal (which would indicate a strong relationship), the adherence rates vary freely at different levels of α. That is, the recognition validity is not indicative of how often the participants followed the heuristic. The correlation between participants' αs and their adherence rate is small ($r = -.19$, $p = .24$), a result that disagrees with both the threshold and the matching hypotheses.[8]
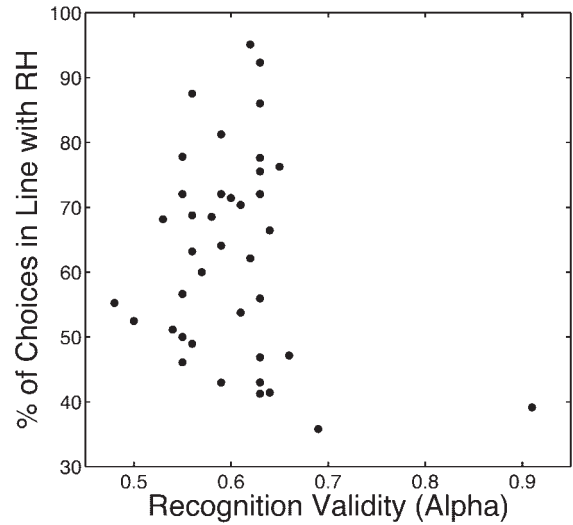
Figure 4. Adherence to the recognition heuristic (RH) as a function of recognition validity.

Finally, according to the suspension hypothesis, object-specific knowledge in conflict with the recognition heuristic can prompt the user to suspend its use temporarily. Assuming that objects differ in the degree to which they are associated with such knowledge, the hypothesis implies varied adherence rates across objects. To investigate this possibility, we calculated for each infection (averaged across participants) the proportion of cases in which the infection was inferred to be the more frequent one, provided that it was recognized and paired with an unrecognized infection. Figure 5 plots these proportions, separately for each infection (averaged across participants). Indeed, there were large differences between the infections. Some, such as gastroenteritis (.97) and viral hepatitis (.91), were almost invariably chosen over unrecognized ones (when the former were recognized). In contrast, infections such as cholera (.30) and leprosy (.15) were mostly inferred to be the less frequent ones. As Table 1 and Figure 5 show, adherence rates are by no means closely lined up with recognition rates ($r = -.10$, $p = .98$): Commonly recognized infections such as cholera, leprosy, malaria, and diphtheria are not necessarily those that command high adherence rate to the recognition heuristic. What drives people's decisions to distrust recognition? We suspect it is the direct and conclusive knowledge that infections such as cholera and leprosy are virtually extinct in Germany, a possibility that we further explore in Study 2.

To summarize, we investigated three candidate hypotheses underlying the restrained use of the recognition heuristic in an environment in which the heuristic does not promise to be highly successful. Two of the three hypotheses—the threshold and the matching hypotheses—received little support: People did not invariably draw on the heuristic as a function of whether their recognition validities α surpassed a threshold (threshold hypothe-
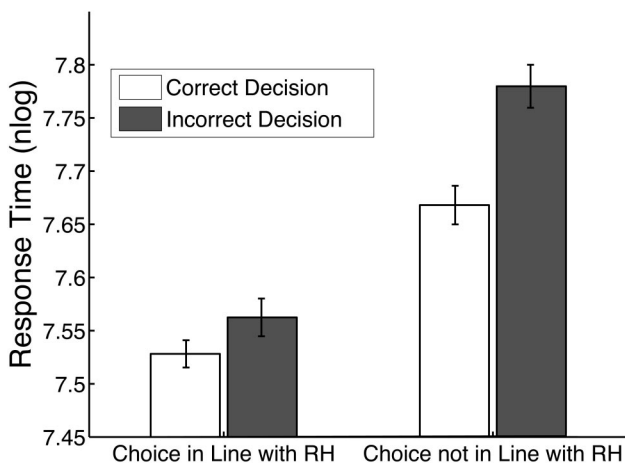

Figure 3. Response times of choices where the recognition heuristic (RH) was applicable as a function of whether or not the choice was in line with the recognition heuristic, and of the accuracy of the choice. Error bars indicate standard errors.

[8] As inspection of Figure 4 reveals, the negative correlation is mainly due to a single participant. When this outlier is excluded, the correlation is $r = .02$ ($p = .91$).
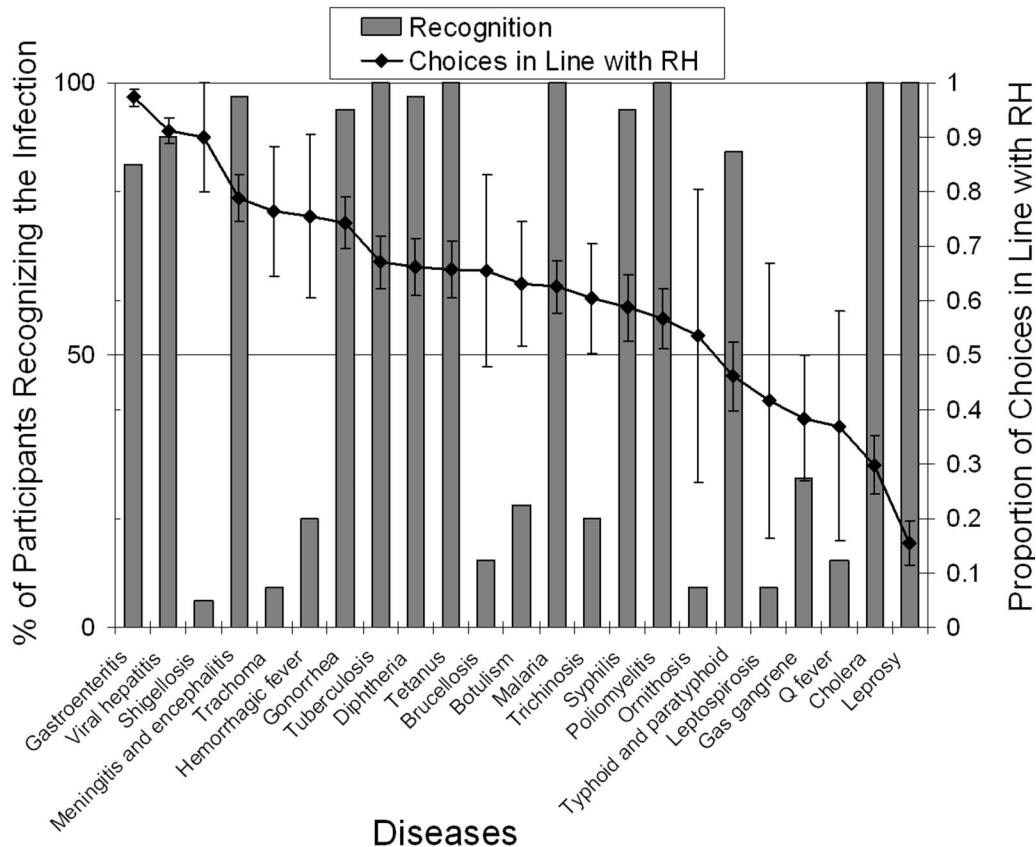
Figure 5. Object-specific recognition and adherence to the recognition heuristic (RH). For each disease, the average (across participants) proportion of choices in line with the predictions of the recognition heuristic (when the disease was recognized and paired with an unrecognized one) is shown. Because only one person recognized it, tularemia is not shown. Error bars indicate standard errors. The columns represent the percentage of participants who recognized the disease.

sis). Similarly, users of the recognition heuristic did not use it with a probability corresponding to their αs. Instead, we observed (a) only a small correlation between individuals' αs and their heuristic adherence rates and (b) enormous variability across infections in terms of participants' reliance on the recognition heuristic. The latter finding suggests that it is object-specific conflicting knowledge that prompts users not to use the heuristic. This finding raises the question of whether relying on this knowledge helped people to boost their inferential accuracy.

Could participants boost their inferential accuracy by temporarily suspending the recognition heuristic? Suspending the recognition heuristic temporarily can improve a person's accuracy (compared with the person's individual α, representing the proportion of correct choices he or she would achieve by invariably using the recognition heuristic whenever applicable). This boost in accuracy, however, will occur only if a person's additional knowledge exceeds the accuracy of his or her recognition knowledge. Did such a boost occur?

We tested this possibility as follows: We first turned to the general question of whether people can discriminate at all between cases in which the heuristic arrives at correct inferences and cases in which the inferences are incorrect. An

analysis based on *signal detection theory* (see the Appendix for the rationale and details of the analysis) yielded that participants were indeed able to distinguish—although not perfectly—between cases in which recognition would have been an invalid piece of information and those in which it would prove valid. But did this ability actually translate into a higher accuracy? For each participant, we calculated the actual accuracy among all items in which the heuristic was applicable. Then, we compared this value with the participant's α (the level of accuracy if he or she had invariably applied the heuristic). Compared with their αs, 24 of 40 participants (60%) managed to boost their accuracy (among the cases in which the recognition heuristic was applicable) by occasionally suspending the recognition heuristic. The accuracy of 16 participants worsened. On average, there was no increase in accuracy: Across all participants, the recognition heuristic would have scored 60.3% ($SD = 6.7$) correct. In comparison, the empirical percentage correct was 60.9% ($SD = 7.4$), a nonsignificant difference: paired-samples t test, $t(39) = 0.39$, $p = .70$. In other words, by temporarily suspending the recognition heuristic, people did not succeed in increasing their inferential accuracy beyond the level attainable if they had invariably used the heuristic.

Summary

In the first study, we tested Predictions 1 and 3. Consistent with Prediction 1, we observed markedly shorter response times for recognition-based inferences: Inferences that were in line with the recognition heuristic proved to require substantially less response time than those conflicting with it (see Volz et al., 2005, for similar results). This finding is consistent with the notion of recognition's retrieval primacy. In contrast with other knowledge, recognition information arrives first on the mental stage and thus has a competitive edge over other pieces of information. Yet, people appear to frequently overrule recognition information in an environment in which there is little to no relationship between recognition and the criterion. Indeed, we found that in such an environment, the use of the recognition heuristic was restrained. Compared with the typically very high adherence rates for the recognition heuristic, we observed an average rate of about 62%. Of three candidate hypotheses concerning how a restrained use of the recognition heuristic is implemented, the suspension hypothesis obtained the strongest support (Prediction 3c). Specifically, people appear to decide case by case whether they will obey the recognition heuristic. Moreover, these decisions are not made arbitrarily but demonstrate some ability to discriminate between cases in which the recognition heuristic would have yielded correct judgments and cases in which the recognition heuristic would have led astray. This ability, however, does not result in a performance boost because the level of accuracy in cases in which the heuristic was set aside does not exceed α.

Study 2: Does Time Pressure Increase Adherence to the Recognition Heuristic?

In Study 1, we found evidence supportive of the notion that recognition has a retrieval primacy and that the decision to set aside recognition information requires extra time. On the basis of this evidence, we now turn to Prediction 2: Bounds on the available response time will increase reliance on the recognition heuristic and will result in a higher rate of inferences consistent with it. Study 2 tests this prediction. In addition, by manipulating the time available for an inference, we address a potential objection to our conclusions in the previous study. In Study 1, participants decided whether to respond swiftly or slowly. Response times, however, can be fast or slow for a number of reasons, including the frequency of the item words in natural language (e.g., Balota & Chumbley, 1984; Scarborough, Cortese, & Scarborough, 1977) or the sheer length of the words. As a consequence, the observed differences in response time in Study 1 could be due to factors other than use of the recognition heuristic. To address this objection, in Study 2, we forced participants to respond swiftly, thus reducing the possible impact of the type of infection (i.e., the characteristics of the infection's name). In addition, we controlled in the analysis for the possible impact of infection type on response time.

Finally, Study 2 further investigates the restrained use of the recognition heuristic in a "hostile" environment. In Study 1, we observed that participants temporarily set aside reliance on recognition. Across infections, such suspension was not distributed evenly but was more pronounced for some infections than for others (see Figure 5). We now explore what kind of knowledge triggers the suspension of the recognition heuristic. Consistent with the results of Oppenheimer (2003), one possibility is the presence of *direct* and *conclusive* knowledge of the incidence rate of a recognized infection that conflicts with recognition information. For instance, a person may remember that cholera has been virtually eliminated (in Germany). This knowledge suffices for the person to conclude that cholera cannot be more frequent and is likely to be less frequent than any other infection, irrespective of whether it is recognized. In general, we suggest that direct knowledge on the criterion variable will overrule recognition information if an LMM can be constructed (see Gigerenzer et al., 1991). An LMM can rest on (a) nonoverlapping criterion intervals and (b) precise figures (ranks) that in combination with elementary logical operations can compensate for missing knowledge (e.g., a particular infection is known to be the rarest infection, thus by extension any other infection is more frequent). It is worth pointing out that by measuring the availability of an LMM independently of the use or nonuse of the recognition heuristic, we can empirically test this view and either refute the suspension hypothesis or accumulate more converging evidence.

Method

Participants and design. Sixty students (none of whom had taken part in Study 1) from Free University (Berlin) participated in the study (41 women and 19 men; mean age = 24.6 years), which was conducted at the Max Planck Institute for Human Development. As in Study 1, the participants were presented with 276 pairs of infectious diseases and were asked to choose the one with the higher annual incidence rate. Furthermore, each participant indicated which infections he or she recognized. Half of the participants took this recognition test before the choice task and half after. They received an initial fee of 9 Euros ($11.80 U.S.) and earned 4 cents (5 cents U.S.) for each correct answer and lost 4 cents for each wrong answer.

Material. Participants responded to the same 276 infection pairs used in Study 1. In addition, they classified each infection in one of the following six frequency categories: < 1–9, 10–99, 100–999, 1,000–9,999, 10,000–99,999, and > 100,000.

Procedure. Participants read the same introductory text as in Study 1 (see previous Method section), after which they were presented with pairs of infections. Time pressure in this choice task was realized as follows (Figure 6): The pairs of infections were presented sequentially on a computer screen in 12 blocks. Each presentation was preceded by an acoustic signal (Tone 1, 10 ms in length), followed by a second signal (900 ms later) that coincided with the presentation of a small fixation cross in the middle of the screen. Again 900 ms later, the cross disappeared, and a third signal followed, accompanied by a pair of infections (left and right from the location of the fixation cross). The pair remained on the screen for 700 ms before disappearing. Participants indicated their response by pressing one of two keys on the keyboard. They were instructed to respond as quickly and as accurately as possible, but not later than a fourth *imaginary* signal, 900 ms after the third tone and the onset of the stimulus presentation (i.e., the signals followed each other in equally spaced intervals; see Figure 6). The reason for using an imaginary signal was to avoid interference of the signal indicating the response deadline with the processing of the stimulus pair (this is a procedure used in research on the lexical decision task; see, e.g., Wagenmakers, Zeelenberg, Steyvers, Shiffrin, & Raaijmakers, 2004). If a response was markedly delayed (i.e., > 1,200 ms after the presentation of the stimulus pair), the message "too late" would appear on the screen, accompanied by an aversive tone. A delayed response reduced the participant's income by 4 cents (5 cents U.S.). In the recognition task, participants saw the names of the 24 infections one at a time (in random order) on the computer screen. They were asked to decide whether they had heard of the infection and to
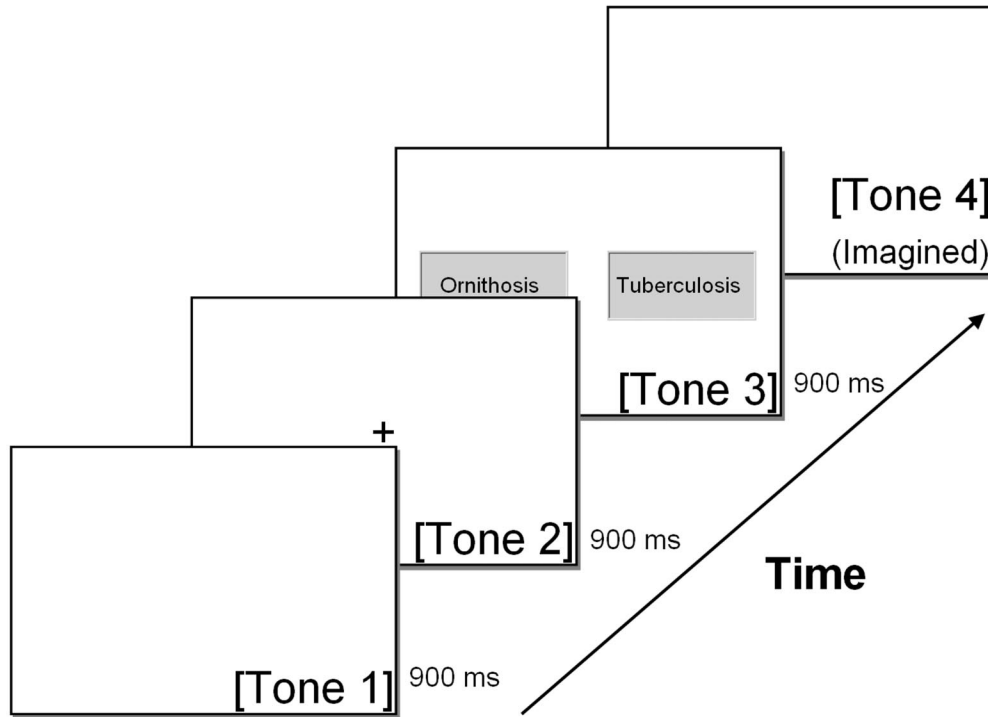
Figure 6. Induction of time pressure in Study 2.

express their positive or negative answer by pressing one of two keys. At the close of the experiment, every participant was asked to classify each infection in one of six frequency categories and to determine whether this judgment was made on the basis of certain knowledge of the criterion variable.

To acquaint participants with the procedure in the choice task, we gave them 10 practice trials. Each practice trial consisted of a pair of arrows (">" and "<", randomly ordered). The task was to indicate within the time limit whether the ">" arrow was shown on the left or right side of the screen. In a second block of 10 practice trials, arrows were replaced by the names of infections, randomly drawn from the pool of infections (and used again in the main choice task).

Results

We first describe the obtained inferences and recognition judgments in more detail. On average, participants scored 58.8% correct ($SD = 4.9$). The cap on response time resulted in somewhat fewer accurate choices, as a comparison with the average score in Study 1 shows, $t(98) = 1.99$, $p = .049$, $d = .41$. On average, participants recognized 61.0% ($SD = 12.6$, range = 41.6%–100%) of the infections (see Table 1). As in Study 1, the frequency of recognized infections was not affected significantly when the recognition task succeeded the choice task (in fact, it was even slightly lower), $t(58) = -1.07$, $p = .29$. The recognition heuristic was applicable in 46.4% ($SD = 11.8$) of the pairs. A student of veterinary medicine recognized all 24 infections, thus rendering the application of the heuristic impossible. Therefore, the recognition validity α was calculated for only 59 participants. The average α was .62 ($SD = .10$), echoing the value obtained in Study 1 (.60). The knowledge validity β, however, was substantially lower than in Study 1: Ms = .62 versus .66, $t(98) = -3.18$, $p =$

.002. It appears that, under time pressure, participants' ability to retrieve additional knowledge was compromised, thus giving way to more guessing responses when both infections were recognized.

Did time pressure increase adherence to the recognition heuristic (Prediction 2)? Consistent with Prediction 2, the proportion of choices in accordance with the recognition heuristic rose under time pressure. Bearing in mind the potential problems with cross-experimental comparisons, the mean proportion of inferences agreeing with the heuristic was 69.2% ($SD = 10.7$, range = 41.4%–90.0%), compared with 62.1% in Study 1, $t(63.5) = 2.5$, $p = .02$, $d = 0.55$. Moreover, the variance in adherence rate (across participants) was smaller in Study 2 than in Study 1, $F(1, 97) = 10.6$, $p = .02$. Note that this increase in the use of the recognition heuristic is not trivial. Time pressure could simply have provoked more guessing. In that case, the proportion of inferences agreeing with the heuristic would have dropped rather than risen. Instead it appears as if time pressure both fostered the use of the recognition heuristic and preempted the retrieval of more knowledge (thus attenuating β).

As Figure 7 shows, the increase in adherence to the recognition heuristic was also manifest on the level of individual infections. For 16 of the 23 infections (70%; as in Study 1, tularemia was not included), more choices agreed with the recognition heuristic than in Study 1 (see also Table 1). In addition, five of the six diseases for which the adherence rate dropped were among the seven diseases with the highest adherence rate in Study 1, thus suggesting a regression effect.

Were inferences in accordance with the recognition heuristic made faster (Prediction 1)? Study 2 also provides another test of Prediction 1. Specifically, we can examine whether within the
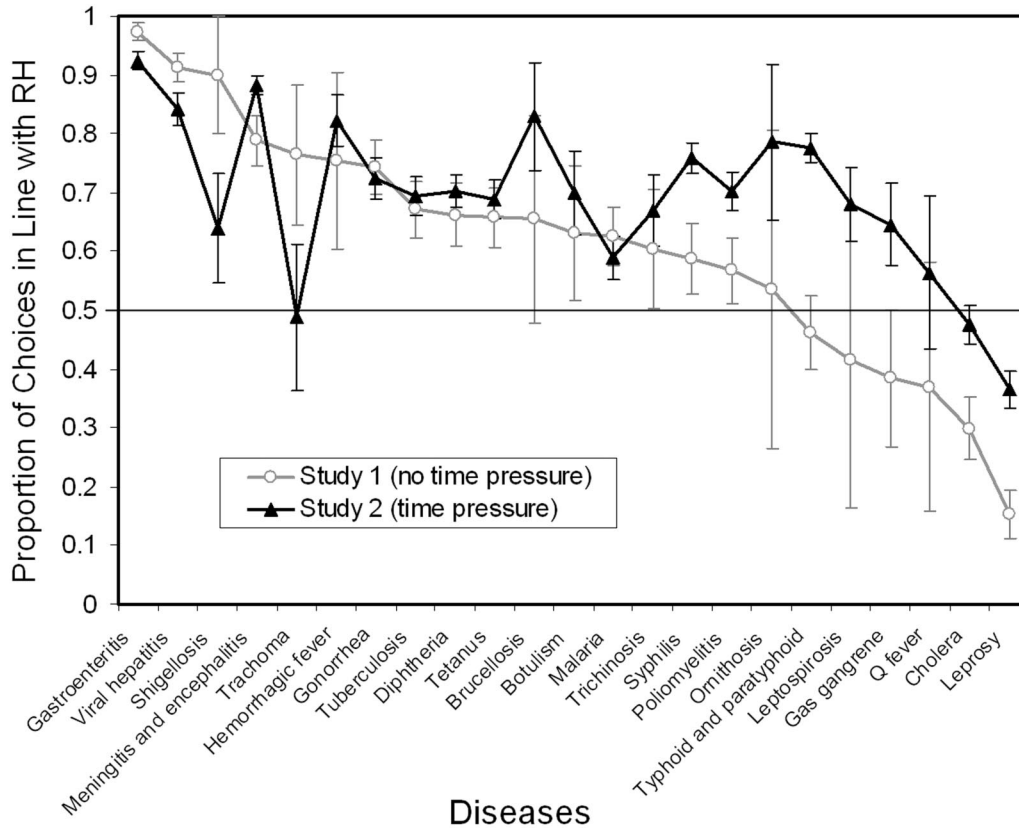
Figure 7. Object-specific adherence to the recognition heuristic (RH) for Studies 1 and 2 (cf. Figure 5). Tularemia, recognized by only one participant each in the two studies, is not shown. Error bars indicate standard errors.

limited response-time window the inferences agreeing with the recognition heuristic declined as a function of time. Such an outcome would support Prediction 1, according to which inferences in line with the recognition heuristic are made faster than those that conflict with it. We divided the response-time window into eight bins, starting with 400–499 ms and ending with responses that lasted longer than 1,100 ms. (Because few responses took less than 400 ms, we omitted them from the analysis.) We then analyzed, for each bin and each infection (for the 15 infections for which there were at least 100 choices within each bin; note that again the choices rather than the participants were taken as the unit of analysis), the proportion of choices in accordance with the recognition heuristic. Figure 8 shows the mean proportions (across infections, thus giving each infection the same weight) in line with the recognition heuristic as a function of response time. Note that as the proportions were calculated across infections, we control for the possibility that the different time bins contained different amounts of choices for the different infections, which could confound the influence of type of infection and response time. Proportions were above 70% for the early bins (i.e., 400–700 ms bins). For later bins, however, the mean proportion dropped rapidly. Consistent with Prediction 1, the more time a response took, the less likely it was to be consistent with the recognition heuristic.

Did conclusive and conflicting criterion knowledge trigger the heuristic's suspension? As Figure 7 shows, choices involving leprosy and cholera (when recognized and paired with an unrec-

ognized disease) resulted in the lowest proportion of recognition adherence in Studies 1 and 2. Why was that? One possibility is that people assumed the diseases to be the least frequent ones. If so, any other disease (even if not recognized) can be inferred to be more frequent than either of the two. Consistent with this view, we found that both infections produced lower frequency estimates than any other infection (see Table 1): The median estimate of their annual incidences was 5.[9] In addition, both infections were those for which the highest proportions of participants (30% and 31.7%, respectively; see Table 1) indicated that they had direct knowledge of incidence rates. These findings suggest that direct and conclusive criterion knowledge for the recognized infection—for instance, knowing that it is virtually extinct—appears to trigger the suspension of the recognition heuristic.

To assess more generally how criterion knowledge impinges on the likelihood of a recognized object being chosen, we reanalyzed people's choices when one infection was recognized but not the other. We focused on those cases in which recognition and criterion knowledge conflicted, specifically those in which the frequency estimate for the recognized infection was *conclusively*

---

[9] We computed these values by replacing each of the six frequency categories (see *Method* section) with the midpoints of each category. For instance, the first category ranging from 1 to 9 was replaced by the value 5. The last category "> 100,000" was replaced by the value 200,000.
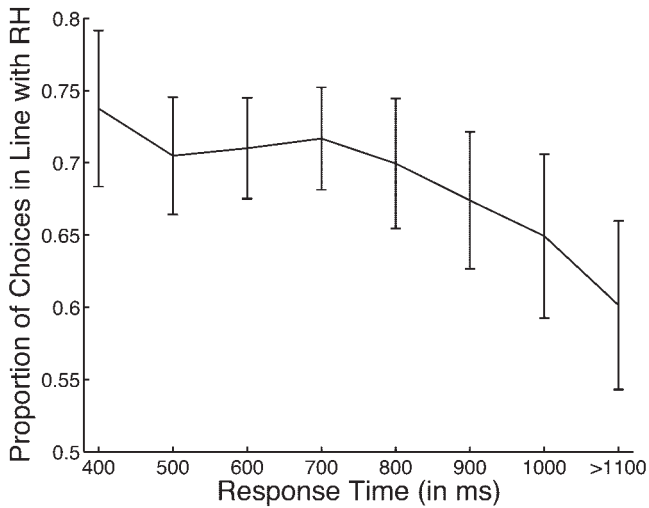
Figure 8. Proportion of choices following the recognition heuristic (RH) as a function of processing time. Over time, there is a decrease in the proportion of choices following the heuristic. The 15 diseases included were gastroenteritis, viral hepatitis, tuberculosis, meningitis and encephalitis, gonorrhea, syphilis, malaria, typhoid and paratyphoid, gas gangrene, hemorrhagic fever, tetanus, leprosy, cholera, diphtheria, and poliomyelitis. The number of choices in the eight bins (400–>1,100) was 141; 747; 1,639; 2,076; 1,354; 752; 260; and 267, respectively.

lower than that for the unrecognized infection. Criterion knowledge was treated as conclusively lower if (a) the estimate for the recognized infection in a pair was lower than the estimate for the unrecognized one by at least two category bins (e.g., the recognized infection was assigned to frequency category "2," the unrecognized one to "4," corresponding to the frequency ranges "10–99" and "1,000–9,999," respectively),[10] and (b) the recognized infection for which a person indicated having direct criterion knowledge was assigned the lowest possible frequency category (i.e., "1") by that person. As pointed out above, both conditions may give rise to an LMM (Gigerenzer et al., 1991), thus rendering reliance on probabilistic cues such as recognition unnecessary. Collapsing across participants, 866 cases (out of 16,560, or 5.2%) met one of the two criteria. The proportion of choices of the recognized infection, averaged across participants for whom there was at least one such case, was below chance level, namely, 45.7%. In addition, the mean proportion of choices of the recognized infection when recognition and criterion knowledge converged was considerably higher, 86.4%, $t(47) = 9.9$, $p = .001$. (Criterion knowledge converging with recognition was defined as cases in which the frequency estimate for a recognized disease was higher than the frequency estimate for the unrecognized infection by at least two category bins; for instance, the recognized infection was assigned to frequency category "4," the unrecognized one to "2," etc.). These results suggest that if an LMM can be constructed, it, rather than the recognition heuristic, guides people's choices.

This conclusion was also corroborated in a reanalysis of participants' choices in Study 1. For this analysis, we took advantage of the median estimates of the diseases' incidence rates obtained in Study 2. Specifically, we focused on those 167 critical pairs (of

11,040, or 1.5%) in Study 1 that contained one unrecognized and one recognized infection, and in which the frequency estimate (from Study 2) for the recognized infection was lower than for the unrecognized one by at least two category bins. The proportion of choices of the recognized infection was 19.1% (again, across all participants where there was at least one critical pair), around 67 percentage points lower than the proportion (of the same participants) in the cases in which recognition and criterion knowledge converged (86.6%), $t(31) = -11.6$, $p = .001$.

Summary

Consistent with Prediction 2, the mean proportion of inferences in accordance with the recognition heuristic increased under time pressure. That is, the competitive edge that recognition information enjoys over other knowledge—its retrieval primacy—translates into more judgments in accordance with the heuristic when people are pressed for time. We also found additional evidence in support of Prediction 1: The longer participants took to make an inference, the lower the proportion of choices in line with the recognition heuristic. Finally, we observed that conclusive and conflicting criterion knowledge appears to be a key condition for the suspension of the recognition heuristic.

General Discussion

When Goldstein and Gigerenzer (2002) proposed the recognition heuristic, they treated recognition as the ability to discriminate between the "novel and the previously experienced" (p. 77). Their intuition was that in many situations an initial sense of recognition (or lack thereof) suffices to make this discrimination. The frugal recognition heuristic does not require additional information such as in which context one encountered the object or what other knowledge about the recognized object one can marshal. Moreover, Goldstein and Gigerenzer assumed that recognition gives rise to noncompensatory inferences: If one object is recognized and the other is not, then the inference can be locked in. Because search for information is then terminated, no other—conflicting—cue information about the recognized object can reverse the judgment suggested by recognition, simply because it is not retrieved. However, this thesis of the recognition heuristic as a strictly noncompensatory strategy has been challenged. Newell and Shanks's (2004) results clearly demonstrate that judgments based on *induced* recognition are reversed when other cues are available that conflict with recognition and when their validity is known to exceed that of recognition. In their view, recognition is a cue as any other.

We aimed to demonstrate that recognition is not like any other cue. To this end, we linked research on the heuristic with research on recognition memory. On the basis of the distinction between a global sense of familiarity and recollection, we proposed that mere

---

[10] It is noteworthy that participants did not consistently give extremely low-frequency estimates for unrecognized diseases. The mean estimated frequency (based on the midpoints of each category) for unrecognized infections was 2,378.0 ($SD = 5,771.4$), which was significantly different from the lowest frequency category, $t(57) = 3.13$, $p = .003$.

recognition is already available while other probabilistic cues are still waiting in the wings. It is retrieved with little to no cognitive effort, whereas other knowledge needs to be searched for. These properties represent what we have termed recognition's retrieval primacy. Based on this notion, we have derived three predictions, and the evidence we have obtained supports them.

Specifically, we found in Studies 1 and 2 that inferences in accordance with the recognition heuristic were made faster than inferences inconsistent with it. In addition, reliance on the recognition heuristic increased when inferences had to be made under time pressure. Finally, we observed that in an environment in which recognition and criterion were not strongly correlated, the recognition heuristic was not as frequently used as had been observed in environments in which there is a strong correlation. Although there are likely to be others (see below), one key factor that triggers the temporary suspension of the use of the heuristic, all other things being equal, seems to be the presence of certain and conclusive knowledge that the recognized object has a low criterion value. In what follows, we discuss the implications of our results. We first turn to a brief review of the conditions that trigger the use and nonuse of the heuristic.

Under What Conditions Do People Use the Recognition Heuristic?

Newell (2005) criticized the fast and frugal heuristics program (Gigerenzer et al., 1999) for having failed to "establish . . . [the] boundary conditions on the adaptive toolbox framework. Without such conditions, it is impossible to evaluate the adequacy of the proposed models of the decision processes" (p. 13). There are now a number of studies on the use of some of the tools of the adaptive toolbox, in particular the Take The Best heuristic (e.g., Bröder, 2000; Bröder & Schiffer, 2003; Newell & Shanks, 2004; Newell, Weston, & Shanks, 2003; Rieskamp & Hoffrage, 1999) and the recognition heuristic. Thus, we can now draw on an—admittedly preliminary—list of conditions that foster and hamper the use of the recognition heuristic.

Recognition validity. The recognition heuristic is useful when there is a strong correlation—in either direction—between recognition and criterion. But attending rather than ignoring recognition can prove helpful, as Goldstein and Gigerenzer (2002) suggested, in any domain in which the recognition validity is higher than chance ($\alpha > .5$). In the current domain of infectious diseases, the recognition validity proved indeed higher than chance ($\alpha = .60$ and .62 in Studies 1 and 2, respectively), and the observed adherence rates, although modest, suggest that recognition was not ignored. Nevertheless, the adherence rate Goldstein and Gigerenzer (2002) observed in the city-size environment (90% in their Study 1) was substantially higher than in the current infection environment (62.1% and 69.2% for our Studies 1 and 2, respectively). Although Goldstein and Gigerenzer did not report the recognition validities of their participants directly, we can take advantage of the recognition correlation (expressed as the rank correlation between the number of participants recognizing a city and its population) they calculated for their city-size environment (see Goldstein & Gigerenzer, 2002, p. 86). Compared with the correlations we obtained ($r_s = .01$ and .03, for Studies 1 and 2,

respectively), the correlation in their study was indeed much higher ($r_s = .66$). One interpretation of these parallel differences is that adherence to the recognition heuristic is at least partly contingent on the correlation between recognition and the criterion in a domain. Such a dependency (which was also observed by Pohl, 2006) may reflect the user's adaptive and ecologically rational use of heuristics (Gigerenzer et al., 1999). Indeed, should one expect an adaptive user of the recognition heuristic to rely on the heuristic to the same extent, irrespective of whether recognition validity is .51 or 1?

However, one should not overstate the degree to which the use of the recognition heuristic may be attuned to α. Individuals' use of the heuristic typically does not, or only moderately, depend on their αs. In addition, such adaptive use of the recognition heuristic is restrained by the accessibility of other knowledge. That is, even in environments with low recognition validity, recognition may often be the only accessible information. To illustrate, in Pachur and Biele's (in press) study of forecasts of the outcome of soccer games, laypeople appeared to rely almost exclusively on recognition in spite of a medium α of .7 and the existence of more valid cues (e.g., team rankings, recent performance). However, these cues are typically available only to experts.

Conflicting knowledge. There are at least three kinds of knowledge that may lead to the suspension of the recognition heuristic: (a) probabilistic cues with validities larger than α (Newell & Shanks, 2004); (b) source knowledge (e.g., the object, say Chernobyl, is known to be recognized for reasons completely unrelated to its size; see Oppenheimer, 2003); and (c) conclusive criterion knowledge (see Study 2; Richter & Späth, 2006, Experiment 1), allowing people to construct an LMM (Gigerenzer et al., 1991). Suspension of the recognition heuristic is, perhaps, least surprising in the case of conclusive criterion knowledge. The very point of the heuristics in the adaptive toolbox (such as the recognition heuristic) is to infer, on the basis of probabilistic cues, an unknown criterion. However, if the criterion is either known or can be deduced, probabilistic inferences will become superfluous (see Gigerenzer et al., 1991). To what extent cues with validities larger than α override recognition information is a bone of contention; we describe our view on this debate in the next section.

Time pressure. Time pressure is conducive to noncompensatory processing (e.g., Dhar & Nowlis, 1999; Payne et al., 1993; Rieskamp & Hoffrage, 1999; Svenson, Edland, & Slovic, 1990; Zakay, 1985; for an overview, see Edland & Svenson, 1993). As we found in Study 2, a cap on response time increased adherence to the recognition heuristic. This result, however, does not simply echo the frequent observation that under time pressure people appear to pay increased attention to the more important attributes in a decision context (e.g., Ben Zur & Breznitz, 1981; Böckenholt & Kroeger, 1993; Kerstholt, 1995; Payne, Bettman, & Johnson, 1988; Wallsten & Barton, 1982). In our view, the reason people make more use of the recognition heuristic under a limited response-time budget is that the retrieval of recognition information precedes that of other pieces of information and requires little to no cognitive effort. Because of these properties, we also suspect that not only time pressure but also, for instance, attending to a second task while performing the choice task would increase adherence to the recognition heuristic.

Is Subjective Recognition a Compensatory or a Noncompensatory Cue: A False Dichotomy?

Goldstein and Gigerenzer (2002) depicted the recognition heuristic as noncompensatory, thus entailing inferences that cannot be reversed by additional *probabilistic cues*. In addition, they showed that treating recognition as a noncompensatory piece of information pays: Specifically, they demonstrated that the performance of compensatory models such as unit-weight or weighted-additive strategies can suffer if recognition is treated like any other cue (Gigerenzer & Goldstein, 1996, p. 660). Challenging the assumed noncompensatory status, Newell and Shanks (2004; see also Bröder & Eichler, 2006) demonstrated that when recognition validity is low, recognition information no longer dominates people's inferences.

One possible way to reconcile these conflicting views is to elaborate the circumstances under which they have been shown to hold. Take, for example, Newell and Shanks's (2004) studies: Recognition information was overruled when participants knew that recognition was an inferior cue—in fact, the worst of all available cues—when they could attribute their sense of recognition unambiguously to one source (the experiment), and when they were cognizant of the presence of superior cues. It seems fair to conclude that people outside of the psychological laboratory may rarely find themselves in such a state of omniscience. How often do we remember the exact source of our recognition knowledge? How often do we know that recognition knowledge is inferior to any other probabilistic cue? Studies investigating processing fluency suggest that its use in inferential tasks is moderated by whether or not it can be attributed to the experiment. By extension, one may expect that recognition is less likely to be overruled in situations in which source knowledge of recognition is nonexistent or diffuse (Johnson, Hastroudi, & Lindsay, 1993), thus suggesting an unspecific, unbiased source of recognition as well as the natural mediation of the criterion variable (see Figure 1).

In contrast to Newell and Shanks (2004), Goldstein and Gigerenzer (2002) investigated the recognition heuristic using naturally evolved recognition. In testing its noncompensatory status, however, they pitted recognition information against a cue that was not consistently superior to α (the validity of the soccer team cue was 78%, whereas Gigerenzer and Goldstein [1996] estimated the recognition validity to be 80%). Thus, unlike in Newell and Shanks's study, and ignoring all other differences, in Goldstein and Gigerenzer's study, recognition information was not ostensibly inferior to that of objective cue knowledge. Thanks to Goldstein and Gigerenzer's and Newell and Shanks's studies and results, one can now ask: Will naturally evolved recognition be overturned by conflicting probabilistic cues with ostensibly higher validity?

Recent studies by Richter and Späth (2006, Experiments 2 and 3) addressed this question, at least partially. As in Goldstein and Gigerenzer (2002), participants were taught additional relevant cue knowledge about objects that they had learned to recognize outside the experimental setting. In contrast to Goldstein and Gigerenzer, Richter and Späth found that the frequency with which a recognized object was chosen to be the larger one in a subsequent inference task was mediated by additional cues. Specifically, recognized objects were less likely judged to be larger when cue knowledge conflicted with recognition knowledge (compared with when recognition and cue knowledge converged). Further investigation is needed, however, of the extent to which and under which conditions probabilistic cues can overturn recognition. For instance, it is still unclear how naturally evolved cue knowledge (rather than induced cue knowledge) interacts with naturally evolved recognition. Moreover, does the relative standing of recognition knowledge and cue knowledge depend on their respective validities?

Regardless of how these questions will be answered, it is important to keep in mind that, for a couple of reasons, recognition is not like any other cue. First, because of its mnemonic properties, recognition represents immediate, insuppressible, and inexpensive information. Studies 1 and 2 demonstrate the implications of these properties for inferences based on recognition. Second, recognition, if applicable, gives rise to an information asymmetry: Because a person typically has no further knowledge about a nonrecognized object, further search *in memory* would typically yield additional information (if any) only about the recognized object. This information asymmetry, in turn, renders the use of information difficult. Hsee (1996; Hsee, Loewenstein, Blount, & Bazerman, 1999) showed that cue values—in particular, continuous cue values—are often ignored when they prove difficult to evaluate. Lack of a reference point (naturally provided by the other object's value), for instance, renders evaluation tricky. Consider, for example, an American student who is asked to infer which of two German cities, Augsburg or Munich, has more inhabitants. She has never heard of Augsburg but has heard of Munich. She also happens to know that Munich has, say, 500 beer gardens—a quantity that she expects to be positively related to city size. However, how big a number is 500? Lacking a standard of comparison (as the corresponding figure for Augsburg is unknown), the student may ignore this cue altogether and rely on recognition only.

Last but not least, with the ongoing debate over the noncompensatory versus compensatory use of the recognition heuristic, it is worth remembering that one of the most robust observations in the evolving science of heuristics is that different people use different heuristics. In other words, there is no single mental strategy that is consistently used by everyone. As has been shown for other fast and frugal heuristics (e.g., Bröder, 2000, 2003; Newell & Shanks, 2003), there may be differences between how people exploit recognition and lack thereof. The recognition heuristic is one model of this exploitation. Some people may only rely on recognition, regardless of whether other cue knowledge is available. Others may use recognition noncompensatorily if its validity exceeds that of other cues. Still others may combine recognition with other cues into a single judgment. The task ahead is to model such individual differences and their link to the probabilistic structure of the environment (e.g., the validity of recognition and other cues).

Are Retrieval Primacy and Availability the Same Thing?

In our view the answer is no. The recognition and the availability heuristics rest on different mental operations. The availability heuristic retrieves *instances* of the target event (e.g., cases of tuberculosis among one's acquaintances) and then bases its inference on either the ease with which such retrieval could be per-

formed or the number of actually retrieved instances (e.g., Tversky & Kahneman, 1974; see also Hertwig et al., 2005). The recognition heuristic, by contrast, bases its inferences simply on the ability (or lack thereof) to recognize the name of the event category (e.g., tuberculosis). To make the distinction crystal clear: The name of an event category can be recognized even if not a single instance is retrievable.

To illustrate, consider Tversky and Kahneman's (1973) classic study involving the categories of famous and nonfamous names. After previously having studied 20 famous and 20 nonfamous names, participants inferred that there were more famous names than nonfamous ones. This inference conforms to the availability heuristic. The recognition heuristic, however, would be mute on this finding. The names of the categories—"famous names" versus "nonfamous names"—are generic labels and equally recognizable (or unrecognizable). This simple example thus illustrates that the availability and the recognition heuristics do not represent different names for the same process. Rather, the availability heuristic may be one of the candidate mechanisms being activated when recognition fails to discriminate (see also Betsch & Pohl, 2002; Goldstein & Gigerenzer, 2002).

How Is the Recognition Heuristic Suspended, and Is Suspension Successful?

We intentionally investigated the recognition heuristic in a real-world domain in which recognition was only weakly correlated with the criterion (see Figure 1). We turned to this "hostile" environment to increase the likelihood of inferences that differ from those determined by recognition and thus to be able to test Predictions 1–3. We found that people relied less on recognition, compared with a domain with a strong correlation between recognition and criterion. Although we do not yet have a solid understanding of how suspension of the heuristic is implemented, we can exclude some candidate mechanisms. First, users do not appear to employ a threshold strategy that demands suspension if α is below a specific threshold (Prediction 3a). Second, users also do not seem to adjust their reliance on recognition to α directly, as described by the matching hypothesis (Prediction 3b).

At this point, the most promising candidate is the suspension hypothesis (Prediction 3c). When time and cognitive resources are available, recognition is followed by an evaluative step in which people assess such aspects as the availability of conclusive criterion knowledge and, perhaps, the availability of source information. From this view, the use of the recognition heuristic could be understood to be akin to a two-stage process proposed in recent memory models. Such models involve a production stage followed by an evaluation stage in which aspects of the production, such as production efficacy, are interpreted and their relevance for a given cognitive task assessed (e.g., Whittlesea, 1997). Indeed, some recent results of a functional magnetic resonance imaging study of the recognition heuristic suggest that recognition knowledge fed into the heuristic might be subjected to such an evaluative filter (Volz et al., 2005).

An evaluative stage that precedes the use of recognition does not contradict the notion that recognition is immediate, insuppressible, and inexpensive. Our thesis of recognition's retrieval primacy only refers to the production of the recognition judgment. It does not refer to the evaluative filter whose activation is likely to require additional cognitive resources (and about whose precise functions one can presently only speculate). It is also unclear whether the evaluative step is a necessary condition for the use of recognition information. This seems to be Newell and Shanks's (2004) view:

> It is not that an object is recognized and chosen without justification, but that the decision-maker has a reasonable idea of why he or she recognizes the object and makes an inference based on this secondary knowledge. Under such an interpretation it is this secondary knowledge that is the driving force behind the inference, not recognition per se. (p. 933)

If, however, secondary knowledge were indeed necessary to clear recognition, the proportion of guesses would be expected to be larger under time pressure (or cognitive load) than under no time pressure, thus causing a decrease in the proportion of choices in line with recognition. We observed the opposite (Study 2).

How successful is the evaluation of recognition information prior to its use? We found that the decision to temporarily suspend the recognition heuristic may not necessarily increase inferential accuracy. To do so, the validity of the knowledge that comes into play when recognition is dismissed[11] must exceed the recognition validity (for this selected set of items). Only then does the user of the heuristic benefit from thinking twice. This raises two interesting issues. First, in environments with a strong correlation between recognition and criterion, it is plainly difficult to top the recognition validity. Thus, high αs in combination with unfavorable odds of finding even more valid information may foster the noncompensatory use of the recognition strategy in such domains. Second and conversely, a low recognition validity in combination with the better odds of finding more valid information may foster the temporary suspension of the recognition heuristic, a speculation consistent with our results.

Conclusions

The recognition heuristic piggybacks on the complex capacity for recognizing objects for making inferences. It bets on a probabilistic link between recognition and environmental quantities, thus turning partial but systematic ignorance into inferential potency. In addition, recognition precedes the arrival of any other probabilistic cue and exacts little to no cognitive cost. Notwithstanding its exceptional properties, the recognition heuristic is only one player in an ensemble of heuristics residing in the mental toolbox. Therefore, there should be limits to its use and boundary conditions that trigger other tools. In this article, we aimed to describe and model some of these conditions. Doing so is key to understanding a heuristic's psychology. In this sense, the cumulative research on the recognition heuristic—despite its currently conflicting conclusions—promises to turn into an exemplary case study in an evolving science of heuristics.

---

[11] Note that this knowledge is not necessarily equivalent to the knowledge captured by the β parameter. It can also encompass knowledge regarding the source of one's recognition and direct criterion knowledge.

References

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Lucek, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 1, pp. 243–293). San Francisco: Freeman.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica, 47,* 89–104.

Betsch, T., & Pohl, D. (2002). Tversky and Kahneman's availability approach to frequency judgment: A critical analysis. In P. Sedlmeier & T. Betsch (Eds.), *Etc.: Frequency processing and cognition* (pp. 109–119). Oxford, England: Oxford University Press.

Böckenholt, U., & Kroeger, K. (1993). Time pressure effects in multi-attribute binary choices. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 195–214). New York: Plenum Press.

Bröder, A. (2000). Assessing the empirical validity of the "Take The Best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1332–1346.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 611–625.

Bröder, A., & Eichler, A. (2006). The use of recognition and additional cues in inferences from memory. *Acta Psychologica, 121,* 275–284.

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.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review, 3,* 37–60.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory and Cognition, 28,* 923–938.

Dawkins, R. (1989). *The selfish gene.* Oxford, England: Oxford University Press.

Dennett, D. (1996). *Kinds of minds.* New York: Basic Books.

Dhar, R., & Nowlis, S. M. (1999). The effect of time pressure on consumer choice deferral. *Journal of Consumer Research, 25,* 369–384.

Edland, A., & Svenson, O. (1993). Judgment and decision making under time pressure: Studies and findings. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 27–38). New York: Plenum Press.

Elrod, T., Johnson, R. D., & White, J. (2005). A new integrated model of noncompensatory and compensatory decision strategies. *Organizational Behavior and Human Decision Processes, 95,* 1–19.

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

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–669.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, 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: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Gronlund, S. D., & Ratcliff, R. (1989). The time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 846–858.

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., & Todd, P. M. (2003). More is not always better: The benefits of cognitive limits. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 213–231). Chichester, England: Wiley.

Herzog, S. (2005). *The boundedly rational fluency heuristic: How ecologically valid is recognition speed?* Unpublished master's thesis, University of Basel, Basel, Switzerland.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory and Language, 33,* 1–18.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Belmont, CA: Duxberry Press.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67,* 247–257.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluation of options: A review and theoretical analysis. *Psychological Bulletin, 125,* 576–590.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology, 56,* 326–338.

Johnson, M. K., Hastroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Juslin, P., & Olsson, H. (2005). Capacity limitations and the detection of correlations: Comment on Kareev (2000). *Psychological Review, 112,* 256–267.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107,* 397–402.

Kerstholt, J. H. (1995). Decision making in a dynamic situation: The effect of false alarms and time pressure. *Journal of Behavioral Decision Making, 8,* 181–200.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Marewski, J. N., Gaissmaier, W., Dieckmann, A., Schooler, L. J., & Gigerenzer, G. (2005, August). *Ignorance-based reasoning? Applying the recognition heuristic to elections.* Paper presented at the 20th Biennial Conference on Subjective Probability, Utility and Decision Making, Stockholm, Sweden.

Martin, R. A. (1983). *Human brain evolution in an ecological context: Fifty-second James lecture on the evolution of the human brain.* New York: American Museum of Natural History.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 563–582.

Mead, M. (1935). *Sex and temperament in three primitive societies.* New York: Morrow.

Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences, 9,* 11–15.

Newell, B. R., & Fernandez, D. (in press). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Journal of Behavioral Decision Making.*

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision-making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 53–65.

Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision-making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 923–935.

Newell, B. R., Weston, N., & Shanks, D. R. (2003). Empirical tests of a fast and frugal heuristic: Not everyone "takes-the-best." *Organizational Behavior and Human Decision Processes, 91,* 82–96.

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90,* B1–B9.

Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science, 15,* 100–105.

Pachur, T., & Biele, G. (in press). Forecasting by ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica.*

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.

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

Payne, S. J., Richardson, J., & Howes, A. (2000). Strategic use of familiarity in display-based problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1685–1701.

Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making, 19,* 251–271.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology, 21,* 139–155.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333–367.

Reimer, T., & Katsikopoulos, K. V. (2004). The use of recognition in group decision-making. *Cognitive Science, 28,* 1009–1029.

Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 150–162.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics and how can we tell. In G. Gigerenzer, P. M. Todd, & the ABC Research Group, *Simple heuristics that make us smart* (pp. 141–167). New York: Oxford University Press.

Robert Koch Institute. (2001). *Epidemiologisches Bulletin* [Epidemiological Bulletin] (Vol. 20). Berlin: Author.

Scarborough, D. L., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 1–17.

Scarr, S., & Salapatek, P. (1970). Patterns of fear development during infancy. *Merrill-Palmer Quarterly, 16,* 53–90.

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, 45,* 513–523.

Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112,* 610–628.

Serwe, S., & Frings, C. (in press). Who will win Wimbledon 2003? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making.*

Snook, B., & Cullen, R. M. (2006). Recognizing national hockey league greatness with an ignorance-based heuristic. *Canadian Journal of Experimental Psychology, 60,* 33–43.

Svenson, O., Edland, A., & Slovic, P. (1990). Choices between incompletely described alternatives under time stress. *Acta Psychologica, 75,* 153–169.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Volz, K. G., Schooler, L. J., Schubotz, R. I., Raab, M., Gigerenzer, G., & von Cramon, D. Y. (in press). Why you think Milan is larger than Modena: Neural correlates of the recognition heuristic. *Journal of Cognitive Neuroscience.*

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys, 14,* 101–118.

Wagenmakers, E. J. M., Zeelenberg, R., Steyvers, M., Shiffrin, R. M., & Raaijmakers, J. G. W. (2004). Nonword repetition in lexical decision: Support for two opposing processes. *Quarterly Journal of Experimental Psychology, 57A,* 1191–1210.

Wallis, G., & Bülthoff, H. (1999). Learning to recognize objects. *Trends in Cognitive Sciences, 3,* 22–31.

Wallsten, T. S., & Barton, C. (1982). Processing probabilistic multidimensional information for decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 361–384.

Whittlesea, B. W. A. (1997). Production, evaluation, and preservation of experiences: Constructive processing in remembering and performance tasks. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 37, pp. 211–264). San Diego, CA: Academic Press.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Modern Language, 46,* 441–517.

Zakay, D. (1985). Post-decision confidence and conflict experienced in a choice process. *Acta Psychologica, 58,* 75–80.

(Appendix follows)

Appendix

To address the question of whether people are able to discriminate between cases in which the heuristic arrives at correct inferences and those for which the inferences are incorrect, we used signal detection theory (Green & Swets, 1966). This theory describes a decision maker who must choose between two (or more) alternatives on the basis of ambiguous evidence. This uncertain evidence is summarized by a random variable that has a different distribution under each of the alternatives, here correct versus incorrect inferences when the recognition heuristic is used. The evidence distributions typically overlap, thus sometimes evidence is consistent with both alternatives. To render a discrimination between the alternatives possible, the person establishes a decision criterion c that divides the continuous strength of evidence axis into regions associated with each alternative. Applied to the question examined here, if the evidence value associated with the event in question exceeds c, the person will conclude, "Following the recognition heuristic leads to a correct inference." Otherwise he or she will conclude, "Following the recognition heuristic leads to an incorrect inference." The person's conclusions can result in four types of outcomes: hits (use of the recognition heuristic yields a correct inference), correct rejections (suspending it yields a correct inference), misses (suspending it yields an incorrect inference), and false alarms (use of the recognition heuristic yields an incorrect inference).

One measure of a person's ability to distinguish between cases in which the recognition heuristic ought and ought not to be used is the distance between the means of the distributions under the two alternatives. If this sensitivity index, d', is small (i.e., the two distributions overlap considerably), a person's decision to temporarily suspend the recognition heuristic is not likely to be more accurate than chance. Across all participants, the observed mean d' differed significantly from zero ($M = .56$ $SD = .43$), $t(38) = 8.11$, $p = .001$. Because 1 participant had a false-alarm rate of zero, the sensitivity measure d' could be calculated for only 39 participants. The d' measure was highly correlated with the sensitivity measure A' ($M = .67$, $SD = .11$), $r = .98$. The mean hit and false-alarm rates were .70 ($SD = .16$) and .50 ($SD = .21$), respectively. Participants thus exhibited some ability to distinguish between cases in which recognition would have been an invalid piece of information and those in which it would prove valid.