Evolutionary cognitive psychology

Peter M. Todd, Ralph Hertwig, Ulrich Hoffrage

Chapter for the *Handbook of Evolutionary Psychology*, ed. by David Buss

Final Draft—August 20, 2004
Not for distribution, but comments welcome!

Introduction: Selective pressures on cognitive mechanisms

Traditional cognitive psychology, the study of the information processing mechanisms underlying human thought and behavior, is problematic from an evolutionary viewpoint: Humans were not directly selected to process information, nor to store it, learn it, attend to it, represent it—nor even, in fact, to think. All of these capacities, the core topics of cognitive psychology, can be seen as epiphenomena arising over the course of evolution from the need to get the central jobs done: survival and reproduction. Moreover, while the subtasks of those two main goals—finding food, maintaining body temperature, selecting a mate, negotiating status hierarchies, forming cooperative alliances, fending off predators and conspecific competitors, raising offspring, etc.—surely relied on gathering and processing information, meeting the challenges of each of these domains would only have been possible by in each case gathering specific pieces of information and processing it in particular ways. This suggests that to best study the faculties of memory, or attention, or reasoning, we should take a task- and domain-specific approach that focuses on the use of each faculty for a particular evolved function, just the approach exemplified by the other chapters in this handbook.

But there is another tack that a traditional faculty-oriented cognitive psychologist can take when facing our domain-oriented mind. In addition to the selective pressures shaping domain-specific mechanisms, there are also a number of important selective forces operating across domains more widely, such as those arising from the costs of decision time and information search. Much as our separate physiological systems have all been shaped by a common force for energy-processing efficiency, individual psychological information-processing systems may all have been shaped by various common pressures for information-processing efficiencies. These broad pressures can in turn lead to common design features in many cognitive systems, such as decision mechanisms that make choices quickly based on little information. As a consequence, cognitive psychologists studying mental mechanisms from a domain-agnostic perspective can benefit from and contribute to an evolutionary perspective that takes into account both domain-specific as well as broad selective pressures.

In this chapter, we show how a set of broad forces operating on multiple domains can impact on the design of specific cognitive systems. In particular, we first discuss how the costs of gathering information, and of using too much information, can be reduced by decision mechanisms that rely on as little information as possible—or even a lack of information—to come to their choices. Next, we explore how the pressures to use small amounts of appropriate information may have produced particular patterns of forgetting in long-term memory and particular limits of capacity in short-term memory. Finally, we show how selection for being able to think about past sets of events has given us reasoning mechanisms best able to handle information represented as samples or frequencies of experience, rather than as probabilities.

Throughout the chapter, we focus on three topics of central interest to cognitive psychologists—decision making, memory, and representations of information. But at the same time, we also lay out three main theses that will be less familiar to those taking a traditional view of cognition as computation unfettered by external, environmental considerations. These theses are that simple decision mechanisms can work well by fitting environmental constraints; limited memory systems can have adaptive benefits; and experience-based representations of information can enhance decision-making. In more detail, we first illustrate how considering broad selective pressures arising from the constraints of information-gathering in the external world can help us to uncover some of the classes of decision mechanisms that people use. Second, in the context of memory systems shaped by such selective pressures, we demonstrate that an evolutionary perspective stressing both benefits and costs of particular abilities can lead to an appreciation of the positive functional roles of cognitive limitations. And third, we argue that taking into account the selective forces exerted by our patterns of interaction with the environment can help explain why different representations of the same information can interact with our evolved machinery to produce widely varying responses. In this way, while we ignore many of the topics typically covered in cognitive psychology, we aim to sketch out some existing questions that we think an evolution-savvy cognitive psychology should explore. (For other views of evolutionary cognitive psychology, and consideration of further issues such as individual differences, see Kenrick, Sadalla, & Keefe, 1998.)

Decision making: Putting information to use

We begin by considering decision mechanisms, which process perceived and stored information into choices leading to action. Cognitive psychology texts typically begin with perceptual and attentional processes and then work their way through the mind finally to decision making and reasoning. To the extent that perceptual systems have to provide information to a variety of domain-specific mechanisms "downstream," they have been shaped through the intersection of multiple selective forces to operate adaptively in a domain-general manner. Vision is the prime example, where the demands of collecting information for foraging mechanisms, mate-selection mechanisms, navigation mechanisms and the like have melded together to select for a visual system that meets general design criteria such as the ability to detect motion and recognize objects in widely-varying lighting conditions. On the other hand, because decision processes stand close to ultimate expressed behavior, they are also close to the particular functionally-organized selective forces operating on behavior. Thus decision mechanisms may have been strongly affected by individual selective forces to become domain-specific. Nonetheless, there are also broad selection pressures operating across domains that, we propose, have shaped a wide range of decision mechanisms in common directions.

What selective pressures impact on decision mechanisms? Foremost, of course, is selection for making an appropriate decision in the given domain. This domain-specific pressure does not imply the need to make the best possible decision, but rather one that is good enough (a satisficing choice, as Herbert Simon, 1955, put it), and on average better than those of one's competitors, given the costs and benefits involved. Good decisions depend on good information, and the specific requirements of the functional problem along with the specific structure of the relevant environment will determine what information is most useful (e.g., valid for making adaptive choices) and most readily obtained. Analyses of the problem and environment structure for particular domains indicate for instance that cues of facial symmetry are relevant and easy to assess for making inferences about mate quality (see Sugiyama, this volume), and that features indicating the presence of refuge and prospect (lookout) locations underlie good decisions about habitat choice (see Silverman & Choi, this volume).

But along with the obvious benefits of gathering information for making decisions come costs, and attendant selection pressures (Todd, 2001), which cognitive psychologists studying the adaptive nature of inference should carefully attend to as well. First, there is the cost of obtaining the information itself. This cost may be paid in temporal or energetic terms: Searching for information can take time that could be better spent on other activities, and can involve expending other resources (exertion in scouting out a landscape, exchange of goods to find out about a potential social partner). Furthermore, such costs can arise in both external information search in the environment, and internal search in memory (Bröder & Schiffer, 2003).

Second, even if information were free and immediately accessible, there is the cost of actually making worse decisions if too much information is taken into consideration. Because we never face exactly the same situation twice, we must generalize from our past experience to new situations. But because of the uncertain nature of the world, some of the features of earlier situations will just be noise, unconnected to the new decision outcome (did interviewing with red underwear on really get me that job offer?). If we consider too much information, then, we are likely to add noise to our decision process, and <u>overfit</u> when generalizing to new circumstances—that is, make worse decisions than if less information had been considered (Martignon & Hoffrage, 2002).[1]

Given these seemingly opposing selective pressures, to make good choices but to do so using little information, what kind of decision mechanisms could possibly be built by evolution? As it turns out, there is little need for a tradeoff between these costs and benefits—many environments are structured such that little information suffices to make

[1] Note that in some situations these potential costs of using too much information may be outweighed by the benefits that seeking extra information can occasionally bring, either directly in decision-making terms as considered for instance in Error Management Theory—see Haselton and Nettle, this volume, or indirectly in social terms such as being able to justify one's diligence to bosses or clients. Nonetheless, while humans may act as information-hungry "informavores" in some domains (Pirolli & Card, 1999), analyzing the costs and benefits incurred by information seeking should guide us in exploring the cognitive mechanisms used in each case.

appropriate choices, and decision mechanisms that operate in a "fast and frugal" manner can outperform those that seek to process all available information (Gigerenzer, Todd, and the ABC Research Group, 1999; Payne, Bettman, & Johnson, 1993). We now briefly survey some of the types of decision heuristics people use that flourish at the intersection of these selective forces. Together, these heuristics form part of the adaptive toolbox of cognitive mechanisms that humans draw on to make adaptive choices in the environments we face (Todd, 2000).

Decision making using recognition and ignorance

Minimal information use can come about by basing decisions on a lack of knowledge, capitalizing on one's own ignorance as a reflection of the structure of the environment. If there is a choice between multiple alternatives along some criterion, such as which of a set of fruits is good to eat, and if only one of the alternatives is recognized and the others are unknown, then an individual can employ the recognition heuristic to guide decision making: choose the recognized option over the unrecognized ones (Goldstein & Gigerenzer, 1999, 2002). Following this simple heuristic will be adaptive, yielding good choices more often than would random choice, only in particular types of environments—specifically, those in which exposure to different possibilities is positively correlated with their ranking along the decision criterion being used. Thus, in our food choice example, the recognition heuristic will be beneficial because those things that we do not recognize in our environment are more often than not inedible; humans have done a reasonable job of discovering and incorporating edible fruits into our diet. (See Galef, 1987, for a similar rule used by Norway rats.) People successfully use the recognition heuristic in a variety of domains where the bigger, better, or stronger instances are discussed more, and hence more widely known and recognized, than the smaller, worse, weaker ones. Examples include large cities, important or rich individuals and social groups, and winning sports teams (Goldstein & Gigerenzer, 2002). Note that the recognition heuristic, as all heuristics, does not guarantee a correct choice. In appropriately structured environments, its use will on average be beneficial and lead to good decisions without having to seek any further information. But in situations where the cost of mistakes is high—for instance, environments where some fruits are known because of their extreme toxicity rather than their deliciousness—decisions should be based on more than recognition alone (Bullock & Todd, 1999).

Decision making using a single reason

When the options to be selected among are all known, the recognition heuristic can no longer be applied, and further cues must inform one's choice. The traditional approach to rational decision making stipulates that all of the available information should be collected, weighted properly, and combined before choosing. A more frugal approach is to use a stopping rule that terminates the search for information as soon as enough has been gathered to make a decision. In the most parsimonious version, "one-reason decision making" heuristics (Gigerenzer & Goldstein, 1996, 1999) stop looking for cues as soon as the first one is found that differentiates between the options being considered. In this case, information processing follows a simple loop: (1) select a cue dimension and look for the corresponding cue values of each option; (2) compare the options on their values for that cue dimension; (3) if they differ, then stop and choose the option with the cue value indicating a greater value on the choice criterion; (4) if they do not differ, then return to the beginning of this loop (step 1) to look for another cue dimension.

This four-step loop incorporates two of the important building blocks of simple heuristics: a stopping rule (here, stopping after a single cue is found that enables a choice between the options) and a decision rule (here, deciding on the option to which the one cue points). To fully specify a particular heuristic, we must also determine the order in which cue dimensions are "looked for" in step 1—the information search building block. Among the many possible one-reason decision heuristics, Take The Best searches for cues in the order of their ecological validity—which reflects their correlation with the decision criterion. Take The Last looks for cues in the order determined by their past decisiveness, so that the cue that was used for the most recent previous decision is checked first during the next decision. The Minimalist heuristic lacks both memory and knowledge of cue validities and simply selects randomly among those cues currently available (the only knowledge it uses is the direction of the cues, that is, whether objects with higher cue values tend to have higher or lower criterion values, which is also used by the previous two heuristics).

Though they use just one piece of information to make decisions, these simple heuristics can nonetheless be surprisingly accurate. Take The Best for instance performed better on average than multiple regression, which combined all available information weighted in an optimal manner, when generalizing to new portions of 20 different real-world environments (Czerlinski, Gigerenzer, & Goldstein, 1999). Furthermore, Take The Best looked up on average just a third of the available cues before finding the one discriminating cue it used to make its decision. Thus, heuristics employing this type of one-reason decision making can successfully meet the selective demands of accuracy and little information use simultaneously. They do so by matching and exploiting the structure of information in the environment (for instance, Take The Best capitalizes on a non-compensatory, or roughly exponentially decreasing, distribution of the importance of cues), using the world to do some of the work and thereby staying simpler and more robust (resistant to overfitting) themselves. A similar analysis within the world of linear models was undertaken by Dawes and Corrigan (1974), who pointed out that simplicity and robustness appear there too as two sides of the same coin: simply ignoring much of the available information means ignoring much irrelevant information, which can consequently increase the robustness of decisions when generalizing to new situations.[2]

[2] More recently, Chater (1999; Chater & Vitányi, 2003) has proposed that minds are themselves designed to seek the simplest possible explanation of the environmental structure they encounter. This quest for simplicity seems to be another general principle that applies across multiple cognitive domains; Chater and others (listed in Chater & Vitányi, 2003) have shown its relevance to understanding perception, language processing, and higher-level cognition. The implications of this important idea for an evolutionary approach to cognitive psychology still need to be worked out.

People use these fast and frugal algorithms in environments that have the appropriate structure (Rieskamp & Otto, in preparation), and where information is costly or time-consuming to acquire (Rieskamp & Hoffrage, 1999; Bröder, 2000; Newell & Shanks, 2003). Socially and culturally influenced decision making can also be based on a single reason through imitation (e.g. in food choice— Ariely & Levav, 2000), norm following, and employing protected values (e.g., moral codes that admit no compromise, such as never taking an action that results in human death—see Tanner & Medin, in press). And when a single cue does not suffice to determine a unique choice, people still often strive to use as little information as possible, for instance via an elimination heuristic (Tversky, 1972): Only as many successive cues are considered, each being used to eliminate more and more alternatives, as are necessary to reduce the set of remaining possibilities ultimately to a single viable option. For example, if one were to use an elimination process to decide on a place to live from among a set of possible habitats, one could first eliminate all those that are too far from water, then all those remaining that are too high, then those that are too cold in winter, and so on until one acceptable site is left. Such a procedure, while using more than one cue, still is able to produce good decisions very quickly (Payne, Bettman, & Johnson, 1993), and can be applied to other types of inference such as categorization (Berretty, Todd, & Martignon, 1999).

Choosing from a sequence of options

When choice options are not available simultaneously, but rather appear sequentially over an extended period or spatial region, a different type of decision mechanism is needed. Here in addition to limiting information sought about each alternative, there must be a stopping rule for ending the search for alternatives themselves. For instance, mate search requires making a selection from a stream of potential candidates met at different points in time. Classic economic search theory suggests that one should look for a new mate (or anything) until the costs of further search outweigh the benefits that could be gained by leaving the current candidate. But in practice, performing a rational cost-benefit analysis is typically difficult and expensive in terms of the information needed (as well as making a bad impression on a would-be partner). Instead, a satisficing heuristic, as conceived by Simon (1955, 1990), can be adaptive: Set an aspiration level for the selection criterion being used, and search for alternatives until one is found that exceeds that level. (See Hey, 1981, 1982, for other simple and quick heuristic approaches to sequential search.)

But how should the aspiration level be set? In situations where options that are passed by at one point cannot be returned to again later (which is often roughly the case in mate search), an effective approach is to sample the first few options that are encountered without selecting any of them, and use the highest value seen in that sample as the aspiration level for further search. This cutoff rule can perform very well in terms of maximizing the mean value of the option ultimately chosen, even with small initial samples (Dudey & Todd, 2002), and people have been shown to use it in experimental settings (Seale & Rapoport, 1997). However, this strategy ignores the problem that a prospect you desire may reject you—the mutual choice constraint underlying game-theory models of two-sided matching (Roth & Sotomayor, 1990). One way to take mutual choice into account in the mate search context is to set one's aspiration level near one's own anticipated or estimated mate value, and hence direct courtship effort at those prospects similar in mate value and so more likely to reciprocate. Simple learning rules that adjust one's aspiration level up with every sign of serious interest from potential partners and down with every rejection can quickly lead to well-calibrated aspirations of this sort that result in realistic patterns of assortative mate choice (Miller & Todd, 1998; Todd & Miller, 1999; Simão & Todd, 2003). While the non-specific pressure to find a mate or other sequentially-available resource without too much search can make the general class of satisficing aspiration-level mechanisms advantageous, the details of the particular search domain (such as whether sequences of options may rise or fall in quality over time) may further select for particular types of search rules (such as rules with later or earlier stopping thresholds).

Ecological rationality and evolved decision mechanisms

The heuristics described above, by ignoring much of the available information and processing what they do consider in simple ways, typically do not meet the standards of classical rationality, such as full information use and complete combination of probabilities and utilities. Furthermore, heuristic algorithms may produce outcomes that do not always follow rules of logical consistency. For instance, Take The Best can systematically produce intransitivities among sets of three or more choices (Gigerenzer & Goldstein, 1996). However, when used in appropriately-structured environments, whether ancestral or current, these mechanisms can be <u>ecologically</u> rational, meeting the selective demands of making adaptive choices (on average) with limited information and time.

Ecological rationality implies a two-way relationship between simple heuristics and their environments (Todd, Fiddick, & Krauss, 2000). First, the success of simple heuristics is defined with respect to pragmatic goals in a particular environmental context. Second, the success of simple heuristics is enabled by their fit to environmental structure (Hertwig, Hoffrage, & Martignon, 1999; Martignon & Hoffrage, 2002). This marriage of structure and simplicity explains and predicts the counterintuitive situations in which there is no tradeoff between being fast and frugal and being successful.

Furthermore, different environment structures can be exploited by—and hence call for—different heuristics. But matching heuristics to environment structure does not mean that every new environment or problem demands a new heuristic: The simplicity of these mechanisms implies that they can often be used in multiple, similarly-structured domains with just a change in the information they employ (Czerlinski, Gigerenzer, & Goldstein, 1999). Thus an evolution-oriented cognitive psychologist should explore both the range of (possibly domain-general) simple decision mechanisms appropriate to a particular adaptive problem, and the domain-specific cues in the environment that will allow those mechanisms to solve that problem effectively.

Memory: Retrieving and forgetting information

To the extent that decisions are based on information, this information is either accessed immediately from the external environment, or from past experience stored internally in some form of memory. Beginning with the pioneering work of Hermann Ebbinghaus (1885/1964), cognitive psychologists usually focus on three aspects of human memory—its capacity, its accuracy, and its structure (e.g. Tulving & Craik, 2000; Koriat, Goldsmith, & Pansky, 2000)—but pay little attention to how it has been shaped by selective pressures, those costs and benefits arising through its use for particular functions in particular environments. Recently, however, researchers have begun to investigate the relationship between the design of memory systems and how they meet their adaptive functions. In this section, we describe some of the trends toward putting evolutionary thinking into the study of memory.

Memory has "evolved to supply useful, timely information to the organism's decision-making systems" (Klein, Cosmides, Tooby, & Chance, 2002, p. 306). The evolution of memory to serve this function has occurred in the context of a variety of costs, which also shape the design of particular memory systems. Dukas (1999) has articulated a wide range of costs of memory, including (1) maintaining an item once it has been added to long-term memory, (2) keeping it in an adaptable form that enables future updating, (3) growing and feeding the brain tissue needed to store the information, and (4) silencing irrelevant information. But taking into consideration the demands of decision mechanisms outlined earlier, the two main selective pressures acting on memory systems (particularly long-term memory) appear to be, first, to produce quickly the most useful stored information, and second, not to produce too much information.

These pressures, like the ones we focused on for decision mechanisms, are broad and general—applying to memory systems no matter what domains they deal with. One way to meet these pressures would be to store in the first place just that information that will be useful later. Having limited memory capacity can work to restrict initial storage in this way, as we will see later with regard to short-term memory. In the case of long-term memory, Thomas Landauer (1982) has estimated that a mature person has "a functional learned memory content of around a billion bits" (p. 491). This is much less than the data storage capacity of a single hour-long music CD, suggesting that we are indeed storing very little of the raw flow of information that we experience. On the other hand, most of what little we do remember is nonetheless irrelevant to any given decision, so our memory systems must still be designed to retrieve what is appropriate, and not more. How can this be achieved? One way is through the very process that at first glance seems like a failure of the operation of memory: forgetting.

Long-term memory: Forgetting curves and the statistical properties of information use

John R. Anderson (1990) put forward an approach he called the rational analysis of behavior as a method for understanding psychological mechanisms in terms of their functions or goals—equivalent to Marr's (1982) computational level of analysis, and also the level at which evolutionary psychology should be focused (Cosmides & Tooby, 1987). Having in mind a view of evolution as constrained local optimization (or hill-climbing), Anderson set out to assess the explanatory power of the principle that "the cognitive system operates at all times to optimize the adaptation of the behavior of the organism" (1990, p. 28). Anderson and Milson (1989) took this approach to propose that memory should be viewed as an optimizing information retrieval system with a database of stored items from which a subset is returned in response to a query (such as a list of key terms). A system of this sort can make two kinds of errors: It can fail to retrieve the desired piece of information (e.g., failing to recall the location of one's car), thus not meeting the pressure of usefulness. But if the system tried to minimize such errors by simply retrieving everything, it would commit the opposite error: producing irrelevant pieces of information (and thus not meeting the pressure of parsimony), with the concomitant cost of further examining and rejecting what is not useful. To balance these two errors, Anderson and Milson propose, the memory system can use statistics extracted from past experience to predict which memories are likely to be needed soon, and keep those readily retrievable. Consequently, memory performance should reflect the patterns with which environmental stimuli have appeared and will reappear in the environment.

This argument can be illustrated with the famous forgetting curve, first described by Ebbinghaus (1885/1964): Memory performance declines (forgetting increases) with time (or intervening events) rapidly at first and then more slowly as time goes on, characterizable as a power function (Wixted, 1990; Wixted & Ebbesen, 1991, 1997). Combining this prevalent forgetting function with Anderson's rational analysis framework yields the following prediction: To the extent that memory has evolved in response to environmental regularities, the fact that memory performance falls as a function of retention interval implies that the probability of encountering a particular environmental stimulus (e.g., a word) also declines as a power function of how long it has been since it was last encountered. Anderson and Schooler (1991, 2000) analyzed real-world data sets to find out whether the environmental regularities match those observed in human memory. One of their data sets, for example, consisted of words in the headlines of the *New York Times* for a 730 day period, and they assumed that reading a word (e.g., "Qaddafi") represents a query to the human memory data base with the goal of retrieving its meaning.

At any point in time memories ("Qaddafi") vary in how likely they are to be needed. According to the rational analysis framework, the memory system attempts to optimize the information retrieval process by making available those memories that are most likely to be useful. How does it do that? It does so by extrapolating from the past history of use to the probability that a memory is currently be needed — the *need probability* of a particular memory trace. Specifically, Anderson (1990) suggested that memories are considered in order of their need probabilities, and if the need probability of a memory falls below a certain threshold, it will not be retrieved. Consistent with their view that environmental regularities are reflected in human memory, Anderson and Schooler (1991) found that the probability of a word occurring in a headline of the *New York Times* at any given time is a function of its past frequency and recency of occurrence. In other words, the demand for a particular piece of information to be retrieved drops the less frequently it occurred in the past and the greater the period of time that has passed since its last use. This regularity parallels the general form of forgetting that has so often been observed since the days of Ebbinghaus. From this parallel, Anderson and Schooler concluded that human memory is a highly functional system insofar as it systematically renders pieces of information less accessible when they have not been used for a while. This functionality operates across domains as a response to broad selection pressures for maintaining quick access to information likely to be useful in upcoming situations (and conversely not maintaining access to information less likely to be needed).

The functions of forgetting

Uncluttering the mind: William James, in the *Principles of Psychology* (1890), was among the first psychologists who pointed to the important function of forgetting. In fact, he argued that “in the practical use of our intellect, forgetting is as important a function as recollecting” (p. 679). In his view, forgetting is the mechanism that enables selectivity. Selectivity, in turn, he asserted

> is the very keel on which our mental ship is built. […] If we remembered everything, we should on most occasions be as ill off as if we remembered nothing. It would take as long for us to recall a space of time as it took the original time to elapse, and we should never get ahead with our thinking. (James, 1890, p. 680)

More recently, contemporary psychologists have begun to specify some of the adaptive functions of forgetting. Elisabeth and Robert Bjork (1996), for instance, have argued that it is critical to prevent out-of-date information—say, old passwords or where we parked the car yesterday—from interfering with the recall of currently needed information. In their view, the mechanism that erases out-of-date information is retrieval inhibition: information that is rendered irrelevant becomes less retrievable. Schacter (2001) also stressed the adaptive functions of forgetting. He, for instance, suggested that various types of misattribution occur when only the general sense of what happened, the gist, is recalled, while the experience’s specific details are forgotten. Memory for gist, in turn, may be fundamental, for instance, for the ability to generalize and categorize across specific instances and thus to organize the permanent flux of experiences. Take Schacter’s example of the category “bird” for illustration. To develop a coherent notion of “bird”, a person has to learn that superficially diverse instances such as a cardinal and an oriole are both members of the same category.

Boosting heuristic performance: The benefits of forgetting, however, may extend beyond the general advantage of setting aside needless information. Forgetting may also boost the performance of heuristics that exploit partial ignorance, such as the recognition heuristic described earlier (Goldstein & Gigerenzer, 2002). Ignorance can come from not learning about portions of the environment in the first place, or from later forgetting about some earlier encounters. To examine whether human recognition memory forgets at an appropriate rate to promote the use of the recognition heuristic, Schooler and Hertwig (2004) implemented this heuristic within an existing cognitive architecture framework, ACT-R (Anderson & Lebiere, 1998). This cognitive architecture is particularly suited to the present analysis as it offers both a plausible memory framework and a strong ecological foundation inherited from the rational analysis of memory mentioned earlier; specifically, ACT-R learns by strengthening memory records associated with, for instance, the names of foodstuffs, habitats or people according to a function that takes the frequency and recency with which they were encountered in the environment into account. Schooler and Hertwig's simulations suggest that in the context of the recognition heuristic one function of forgetting is to actively maintain the system's ignorance. In other words, in all their simulations they found that the performance of the recognition heuristic indeed benefited from (a medium amount of) forgetting.

Strategic information blockage: In the case of the recognition heuristic, forgetting refers mostly to content in declarative memory. Could even forgetting of parts of one's autobiography be adaptive? The argument that forgetting of seemingly unforgettable experiences, that is, traumatic experiences, can serve important functions has been entertained since the late 19th century. Since the 1980s the notion of repressed memories — in particular of memory for childhood sexual abuse — has received a great deal of academic as well as public attention. We do not review the controversial debates that ensued pertaining to such questions as whether or not recovered memories can occur and how accurately they correspond to actual events (for an excellent review see Silvers, J. Schooler, & Freyd, 2002). Here we are only concerned with one theory of recovered memory in which temporary forgetting (or reduced accessibility) of traumatic events is assumed to be functional.

Betrayal trauma theory proposed by Jennifer Freyd (1996) suggests that the function of amnesia for childhood abuse is to protect the child from the knowledge that a key caregiver may be the sexual perpetrator. In situations involving treacherous acts by a person depended on for survival, a "cognitive information blockage" (Silvers et al., 2002, p. 177) may occur that results in an isolation of knowledge of the event from awareness. In fact, such temporary forgetting may be a prerequisite for maintaining the crucial relationship with the caregiver, and ultimately, for survival. At least two different mechanisms can account for memory impairments for trauma-related information. One is avoidant processing, in which people disengage attention from threatening information and thus fail to even encode it. Another mechanism locates the cause for the information blockage at the retrieval stage, assuming that threatening information is encoded but cannot be retrieved (see McNally, Clancy, & Schacter, 2001). Betrayal trauma theory also yields specific predictions about the factors that will make this type of forgetting most probable—for instance, it predicts that amnesia will be more likely the more dependent the victim is on the perpetrator (e.g., parental vs. nonparental abuse). While the experimental evidence for betrayal trauma theory is preliminary (Silvers et al., 2002) and controversial (see McNally et al., 2001 and DePrince & Freyd, 2004), the theory illustrates how domain-specific forgetting may have unique adaptive functions.

Hindsight bias: Consequence of future-oriented memory

Of course, forgetting is not always beneficial. Take the *hindsight bias* as an example (see Hoffrage & Pohl, 2003). The hindsight bias is the phenomenon that once people know the outcome of an event, they tend to overestimate what could have been anticipated in foresight. This can come about because people do not have perfect memories of all the opinions and judgments they hold in the past. Therefore, if they have to remember in hindsight how likely they thought it was that, for instance, Al Gore would be elected— in light of their knowledge that he lost — they may overestimate their previous doubts. Fischhoff (1982) stressed the potentially harmful consequences of the hindsight bias: "The very outcome knowledge which gives us the feeling that we understand what the past was all about may prevent us from learning anything from it" (p. 343). Following Fischhoff's lead, the hindsight bias has been seen as a severe and systematic bias in memory.

Even here, however, adopting a functional perspective can provide a deeper understanding of the memory illusion. Hoffrage, Hertwig and Gigerenzer (2000) took such a perspective in their model of the hindsight bias. They assumed that being informed about the outcome of an event (e.g., an election) can result in an *updating* of the knowledge that was originally used (e.g., which candidate has more charisma) to try to infer which outcome would occur (e.g., which candidate would win). In addition, their model assumes that if people cannot retrieve their original judgment, they will reconstruct it by going through the same steps of inference that led to the original judgment, that is, people will (re)simulate their original judgment process. In the mean time, however, some of the elusive and missing cue values have been updated in light of the actual outcome. Therefore, the reconstructed judgment may turn out to be closer to the actual outcome than the original judgment was. This model makes the novel prediction—confirmed by Hoffrage et al.'s studies—that feedback on one variable (e.g., election outcome) can lead to systematic changes not only in the recalled prediction, but also in the memory of associated variables (e.g., cues related to election outcomes such as charisma of the person who lost the election). Moreover, this algorithmic process model is specific enough to explain why hindsight bias occurs, does not occur, or is reversed in particular individual responses (see Hoffrage et al., Figure 5).

Is the hindsight bias detrimental? No doubt, if the goal is to veridically reconstruct previously held judgments, preferences, or opinions, then hindsight bias caused by knowledge updating can be a great hindrance to fulfilling this goal. But such a goal is unlikely to be very common or important. The real world, as opposed to the psychological laboratory, is inherently unstable, and the longer the time interval since the last assessment of, say, a foraging location, the more likely that the environment will have changed and the location will no longer have the same value. As Bartlett (1932/1995) put it: "In a world of constantly changing environment, literal recall is extraordinarily unimportant" (p. 204). In other words, in the trade-off between an accurate remembrance of the past and accurate inferences in the future, emphasizing future performance should win. Our memory seems to be designed to do just this, wagering on the future rather than on the past: Simulations by Hertwig, Fanselow and Hoffrage (2003) suggest that at the same time that knowledge updating may be increasing hindsight bias, it can increase the accuracy of future inferences. On this view, the hindsight bias is a cheap price to pay for an adaptive advantage, namely, the timely provision of useful up-to-date knowledge.

Short-term memory: Functional explanation of its bounds

The previous analyses apply to long-term memory. Long-term memory, however, is only one of the components posited within traditional memory architectures, for instance, the Atkinson-Shiffrin model of human memory (Atkinson & Shiffrin, 1968). Another key component is short-term memory. The classic estimate of the capacity of short-term memory is 7±2 chunks (Miller, 1956), and more recent estimates make it even smaller (Cowan, 2001). Why has it evolved to be so limited? Anderson (1990) feels that it is because we are "trapped on some local optimum of evolution" (p. 92), but this does not seem convincing. While greater short-term memory size may have required somewhat increased brain metabolism or other tradeoffs (Dukas, 1999), there do not seem to be inherent constraints that would have ruled out more generous capacities. In the absence of strong constraints, more plausible explanations for why evolution has produced such a modest mental storage capacity draw on functional considerations. One of the most interesting functional explanations has been put forth by Yakov Kareev in a series of papers (Kareev, 1995a,b; 2000; Kareev, Lieberman, & Lev, 1997). Kareev argues that while limitations of working memory capacity force people to rely on small samples of information, these small samples also have a specific advantage: They can enhance inferences of causality by enabling the early detection of covariation between elements in the environment.[3]

Kareev's argument runs as follows. To determine at a given point whether two variables covary (e.g., does this pile of droppings mean a predator is nearby?), one often needs to rely on data sampled from one's environment or from long-term memory (i.e., past observations of the environment), which is then entered into working memory.[4] To the extent that the degree of covariation is derived from the information that is currently in one's working memory, that system's limits imposes an upper bound on the size of the information sample that can be considered at one time. Taking Miller's estimate as a starting point, Kareev et al. (1997; Kareev, 2000) suggested that using samples of around 7 observations of the co-occurrence (or lack thereof) of two events increases the chances for detecting a correlation between them, compared to using a greater number of observations.

[3] Another proposal for a functional benefit of bounded short-term memory is MacGregor's (1987) theoretical argument that it can speed up information retrieval.

[4] Kareev uses the term "working memory" as akin to the earlier concept "short-term memory," but see Baddely, 2000, on the different meanings of the term "working memory."

The reason is that randomly drawn samples of a population of two-variable observations are likely to show a within-sample correlation that differs somewhat from the true population correlation. More interestingly, looking at *small* randomly-drawn data samples (whether from the environment or from long-term memory) increases the likelihood of encountering a sample that indicates a stronger correlation than that of the whole population. To see why, imagine drawing many small samples of two continuous variables, calculating the relationship between them (i.e., Pearson's product-moment correlation), and plotting the distribution of the correlation coefficients thus found. Then (provided that the population correlation is not zero), the resulting distribution will have a characteristic skewed shape, with both the *median* and the *mode* of the distribution more extreme than the corresponding population values. Moreover, the amount of skewedness is a function of the sample size: The smaller the sample, the more skewed the resulting distribution. (See Figure 1.)

[insert here Figure 1 of skewed distributions from Kareev]

In other words, when drawing a small random sample from a population in which a correlation exists, the sample-based correlation estimate is more likely than not to be more extreme than the true correlation found in the population. Thus, a limited working memory can function as an amplifier of correlations, allowing those present in the population to be detected earlier than they would be if working memory, and the sample size necessary to fill working memory, were larger. Consistent with this thesis, Kareev et al. (1997) found that people with smaller working memory capacity detected correlations faster and used them to make correct predictions better than people with larger working memory capacity. This enhanced ability to detect contingencies seems particularly important in domains in which the benefits of discovering a causal connection outweigh the costs of *false alarms*, which also increase in number with smaller sample sizes (a point highlighted by Juslin & Olsson, in press—but see Fiedler & Kareev, 2004, for further considerations). Such domains may be characterized by situations in which missing potential threats would be extremely costly, including for instance learning about the cues associated with the presence of a predator or the signals presaging a stranger's harmful intentions.

Kareev's (2000) analysis suggests that evolution may have designed the capacity of human working memory to correspond to a "window size" that amplifies existing contingencies in the world, thus fostering their early detection. This thesis not only offers a functional explanation for why temporary mental storage capacity is so limited, it also sheds new light on what has been interpreted to be a cognitive or perceptual bias—the *belief in the law of small numbers*. According to Tversky and Kahneman (1971/1982), "people's intuitions about random sampling appear to satisfy the law of small numbers" (1982, p. 25), in that "people view a [small] sample randomly drawn from a population as highly representative, that is, similar to the population in all essential characteristics" (1982, p. 24), rather than understanding that small samples are more likely to deviate further from population statistics (e.g. the mean) than are large samples. Tversky and Kahneman proceeded from this observation to criticize human intuition for being unduly swayed by small samples, foreshadowing the conjectures about the failings of human rationality that dominated cognitive and social psychology in the decades to come (see e.g. Krueger & Funder, in press). Certainly, overreliance on small samples may indeed exact a price in terms of systematic misperceptions of the world—but the important thing to ask from an evolutionary cognitive psychology perspective is how large that price is compared to the potential benefits accruing to their use. Kareev's analysis can be taken as a challenge to the premise that the more veridical the mental representations of the world, the better adapted the organism.[5]

To conclude our discussion of the selective forces acting on memory, we return to Anderson and Milson (1989), who pointed out that "one seldom finds arguments for a theory of memory mechanisms cast in terms of the adaptiveness of these mechanisms" (p. 703). This situation now seems to be slowly changing, not least because of Anderson and colleagues' leading work. Moreover, exploring the adaptiveness of memory necessitates not only asking what memory was designed to do, but also how the design benefits could be achieved in light of the pertinent costs. This combination of a functionalist view with a cost-benefit analysis of particular mechanisms, as often employed in evolutionary cognitive ecology (Dukas, 1998), will move us closer to a thorough understanding of the workings of human memory.

Representation of Information: Modern Practices meet Evolutionary Constraints

In the previous section we discussed some aspects of memory from an evolutionary point of view. But why do we have memory at all? Why should we be able to recall representations of the past? After all, changes in behavior could arise through learning even without the ability to remember independently any aspects of the events that we learned from. Being able to store and retrieve information about what happened in the past, however, lets us process that information further in the light of new information and experience. It also allows us to communicate the information to others (as well as to ourselves at later points in time) and combine it with information from them in turn. Ultimately, recalled information from the past enables us to form expectations about the future which can guide behavior in the present.[6]

Internal memories, our focus in the previous section, are not the only innovation over the course of evolution for representing past events. Paintings of animals in Pleistocene caves, for instance, demonstrate one step in the development of representations that have

---

[5] Referring to scientific models, William Wimsatt, a philosopher of biology, argued that "False models build locally truer theories" (Wimsatt, 1987), because they isolate aspects of our ignorance and allow us to progress. His supposition is that the "creative use of falsehood is one of the best tools the practicing realist-scientist ... has for discovering truths about nature." Our conjecture is that systematically inaccurate mental models of the world can also confer functional benefits to organisms whose aim is not to explain the world but rather to survive and reproduce in it.

[6] See Freyd, 1983, 1990, for a theory of how pressures for shareability of information between and within individuals can, in conjunction with pressures from natural selection on cognitive systems, shape the representations of information that we use.

been used to externalize internal states—here, memories of what the early artists had previously seen outside the cave. During the evolution of culture such external representations were complemented by symbols that became standardized and gradually reached greater and greater levels of abstraction. Ultimately this led to alphabets and number systems that could be used to convey complex information to others, both contemporaries and successors. Parallel to this process, procedures to collect and combine information have been refined over the centuries, finally leading to huge national and international institutions founded to gather and aggregate demographic, social, economic, medical, and other kinds of data (Gigerenzer, Swijtink, Porter, Daston, Beatty, & Krüger, 1989).

As a consequence, the sources of information that could be used as a basis for judgments and decisions have increased over the course of human evolution, from individual experiences (a source that we share with even the lowest animals), through reports from family or group members (a source that social animals have, and that humans have in greatly developed form), to modern statistics (a source that has been added only very recently during our cultural evolution). Does it make a difference, in terms of individual decision making, what form the information takes as a consequence of its source? Adopting an evolutionary point of view, one would hypothesize that the answer is “yes,” because our cognitive systems have been exposed to different forms and sources of information for different amounts of time. In particular, forms that have been created during our most recent cultural development may pose a bigger challenge to our information processing capacities than those to which the human species had much more time to adapt. In this section we present evidence supporting this hypothesis, showing how different types of representations affect decisions first in situations involving risks, and second in the context of Bayesian inference tasks. As in the previous sections, the selection pressures we consider here apply to the use of information representations across a wide range of functions and tasks, and so will have shaped cognitive mechanisms from many domains in similar ways.

Decisions from experience versus decisions from description

Much of everyday decision making can be understood as an act of weighing the costs against the benefits of the uncertain consequences of our choices. Take the decision of whether to engage in short-term mating as an example. Although casual sex has obvious evolutionary benefits (especially for men; see, e.g., Trivers, 1972), it can cause one to contract a sexually transmitted disease, acquire an undesirable social reputation, or suffer violence at the hands of a jealous partner (for other risks of casual sex from an evolutionary perspective, see Buss, 2004). Each of these consequences, whether beneficial or harmful, is uncertain: It might happen, and it might not happen. Choosing to have casual sex is thus like rolling a die, each side of which represents one or more possible consequences of that choice. Only after the die has come to rest, the decision made, and the action taken, will we find out which of the consequences has become reality. Because uncertainty is an integral part of “virtually all decisions,” wrote Goldstein and Weber (1997), “*life is a gamble*” (p. 569).

The metaphor of life as a gamble has exerted a powerful influence on psychological research on decision making under risk, giving rise, for example, to the ubiquitous use of monetary lotteries in laboratory experiments. Studies that employ such lotteries typically provide respondents with a symbolic—usually written—description of the options, for example:

A: Get \$4 with probability .8, \$0 otherwise. or *B*: Get \$3 for sure.

The most prominent descriptive theory of how people decide between such lotteries is *prospect theory* (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Briefly put, prospect theory assumes that the human mind "frames" the outcomes of a decision in terms of gains and losses. Losses are more painful than gains of the same magnitude are pleasurable, but sensitivity to identical decrements (or increments) in value diminishes as the losses (or gains) increase. Prospect theory further posits that, relative to the objective probabilities with which an outcome can be expected to occur, people make choices as if small-probability events receive more weight than they deserve and as if large-probability events receive less weight than they deserve. This assumption can explain why, for instance, most people (80% of participants in Kahneman & Tversky, 1979) are inclined to choose lottery *B* over *A* above: The rare outcome in *A*, receiving \$0, receives more weight than it deserves, reducing the perceived value of *A*.

But are choices between options like *A* and *B* representative of the gambles with which life presents us? Hertwig, Barron, Weber, and Erev (2004, in press) argue that we rarely have complete knowledge of the possible outcomes of our actions and their probabilities. When deciding whether to have a one-night stand, for instance, we do not make a *decision from description*, consulting a written list of the possible consequences and their likelihoods. Instead, we rely on the experience that we (or others) have accumulated over time. Hertwig et al. referred to this kind of choice as a *decision from experience*.

Do people behave differently when they learn about outcomes and probabilities from written descriptions as opposed to experience? To find out, Hertwig et al. (2004, in press) created an experimental environment in which decision makers started out ignorant of the outcomes and the outcome probabilities associated with pairs of lotteries. On each trial, respondents saw two buttons, here denoted *A* and *B,* on a computer screen and were told that each button was associated with a payoff distribution. When they clicked on a button, an outcome (e.g., \$3 if they chose *B* above, or \$0 on 20% of clicks and \$4 on 80% of clicks if they chose *A*) was randomly sampled from its distribution. Respondents could sample from either distribution as many times as they wished. After they stopped sampling, they were asked which lottery they wanted to play for real payoffs.

Hertwig et al. (2004, in press) compared the choices of respondents who received written descriptions of each option (i.e., the amount of money to be gained or lost and the probability of winning or losing it) with those made by respondents who were allowed to sample the possible outcomes freely and repeatedly as described above. Although respondents in the description group and the experience group were given structurally identical problems, the choices they made differed dramatically between the groups. Across six problems, the average absolute difference between the percentage of respondents choosing the option with the higher expected value (e.g., *A* above) in the experience and description groups was 36 percentage points. Moreover, in every problem, this difference was consistent with the assumption that rare events (e.g., $0 in *A*) had more impact than they deserved (given their objective probability) in decisions from description—consistent with prospect theory—but had less impact than they deserved in decisions from experience.

To account for the dramatic difference between decisions from description and decisions from experience, Hertwig and colleagues cited two factors—small samples and a recency effect. First, the experience group tended to rely on small samples of outcomes, which meant that they either did not encounter the rare event or encountered it less frequently than expected on the basis of its objective probability. Second, they paid more attention to recently experienced outcomes. Most people did not encounter rare events in the last few draws from the payoff distribution because of the very rarity of those events. In contrast, having read about the rare events in their written presentation alongside the common events, the description group appeared not to overlook the rare events but rather to overweight them.

The distinction between decisions from description and decisions from experience not only explains people's different risky choices in structurally identical problems but also points to the solution to an intriguing puzzle related to the behavior of bees. Because animals do not share humans' ability to process symbolic representations of risky prospects, all their decisions (for instance, about where to forage) are decisions from experience. In a study of foraging decisions made by bees, Real (1991) observed that "bumblebees underperceive rare events and overperceive common events" (p. 985). To explain why bees' "decision weights" diverge from those observed in humans and captured by prospect theory, Real cited, among other factors, the fact that bees' samples from payoff distributions are truncated due to memory constraints. Although humans and bumblebees do not share a recent evolutionary history, Hertwig et al.'s (2004) results suggest that the decisions of bumblebees should converge with those of humans when humans, like bees, rely on small samples of experience (see also Weber, Shafir ,& Blais, 2004).

The more general implication of the distinction between decisions from description and decisions from experience is that representations that are identical mathematically can be different psychologically. Furthermore, the two sources of information differ not only in form, but also in the length of time that they have exerted a pressure on our cognitive abilities to understand and process them appropriately. Throughout the course of human evolution we have experienced events in our interactions with the environment, but only very recently have we begun to aggregate such information and communicate it in the form of statistical descriptions.[7] Thus, one might speculate that our cognitive strategies

[7] The further questions of how people use non-scientific language (as opposed to statistics) to communicate subjective likelihoods, via words such as "often", "sometimes", and "rarely", and how these words are

for making decisions under risk are more likely tuned to experienced frequencies than to described probabilities. Not only does the research of Hertwig and colleagues just described support this assertion, but also work done in the domain of Bayesian reasoning—the topic to which we now turn.

Inferences based on natural frequency versus probability representations

How should we update our beliefs in light of new evidence? For instance, how should a Pleistocene hunter update his belief regarding the chance of finding prey at a particular location after he has seen some unusual movements in the grass over there? As this example shows, we have been facing the task of updating beliefs for a long time, and there should have been sufficient selective pressure to produce a mechanism able to perform such inferences. At first glance, however, the current empirical results are inconclusive: Whereas research by Gallistel (1990) and Real (1991) suggests that animals are adept at such Bayesian inferences (updating of beliefs in light of new evidence), humans seem to lack this capability: "In his evaluation of evidence, man is apparently not a conservative Bayesian: he is not a Bayesian at all" (Kahneman & Tversky, 1972, p. 450). How can it be that bumblebees are better at making Bayesian inferences than humans?

As in the previous section, the answer lies in the different ways that information can be represented. How did bumblebees and Pleistocene hunter-gatherers encounter the statistical information about environmental features? On a trial-by-trial basis, that is, by sequentially observing cases—which, in the simplest case of dichotomous variables, means observing whether a predictor is present or absent and whether the criterion is present or absent. Adaptive behavior can be based on accurate judgments of the probability of the criterion being present given that the predictor is present. Such judgments can be made by a mechanism that is sensitive to the difference between the number of cases in which predictor and criterion are present and the number of cases in which only the predictor is present (possibly giving more weight to the most recent cases). Experiments with human participants in which cases are sequentially presented, thereby allowing participants to observe the states of the predictor and the criterion over successive trials in a natural fashion, have shown that people are well able to estimate the probability of observing the criterion given the presence of the predictor (Christensen-Szalanski & Beach, 1982).

In contrast, those studies leading to the conclusion that people are <u>not</u> able to reason in a proper Bayesian fashion have presented participants with descriptions given in terms of probabilities. For example, Eddy (1982) presented 100 physicians with the following information:

---

understood by the audience, is a large research area in itself; see for example Hertwig and Gigerenzer (1999), Dhami and Wallsten (in press).

> The probability of breast cancer is 1% for a woman at age forty who participates in routine screening. If a woman has breast cancer, the probability is 80% that she will have a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also have a positive mammography.

Physicians were then asked to imagine a woman in this age group who had a positive mammography in a routine screening, and to state the probability that she actually has breast cancer. Out of those 100 physicians, 95 judged this probability to be between .7 and .8. To obtain the Bayesian solution, which is usually seen as the normatively correct answer, one “simply” has to insert the probabilities into Bayes’ rule (if one were lucky enough to know about it and remember it):

$$p(C|M) = \frac{p(C)p(M|C)}{p(C)p(M|C) + p(-C)p(M|-C)} = \frac{(.01)\,(.80)}{(.01)\,(.80) + (.99)\,(.096)}$$

where C stands for breast cancer, -C stands for no cancer, and M stands for a positive mammography result. The result is actually .07, indicating that probability-based descriptions led most of the decision makers in Eddy’s study widely astray. (See Figure 2, left panel.)

[insert here Figure 2 of probability and frequency solutions to Bayesian problem from Hoffrage]

However, the differences in decision-making performance here do not come down to just a distinction between beneficial experience and detrimental description. By considering what kinds of representations our minds evolved to deal with, Gigerenzer and Hoffrage (1995) created an effective compromise between sequential acquisition of information and descriptions in terms of probabilities: They presented participants with descriptions in which the probabilities were translated into natural frequencies. Natural frequencies result from natural sampling (Kleiter, 1994) in which cases are randomly drawn from a specified reference class. Thus, while participants did not encounter cases one-by-one, they were presented with the aggregate statistics in terms of numbers that arise when an entire sample of individual cases is experienced and counted.

For illustration, the probability information provided by Eddy (1982) when converted into natural frequencies reads as follows: “Out of 10,000 women, 100 have breast cancer. Out of those 100 women with breast cancer, 80 have a positive mammogram. Out of the remaining 9,900 women without breast cancer, 950 nonetheless have a positive mammogram.” Asking for the probability that a woman has breast cancer given a positive mammogram now becomes “How many of those women with a positive mammogram have breast cancer?” This question, calling for an inference that has to be made from information presented in terms of natural frequencies, is much easier to answer. (See Figure 2, right panel.) Gigerenzer and Hoffrage (1995) first replicated the finding that presenting information in probabilities resulted in poor performance: Across

15 tasks, participants reasoned the Bayesian way only 16% of the time. When the information was presented in terms of natural frequencies, this percentage rose to 46%, and the number of answers that were close to Bayesian rose greatly as well. Similar results were obtained with physicians (Hoffrage & Gigerenzer, 1998), medical students (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000), and lawyers (Lindsey, Hertwig, & Gigerenzer, 2003).

Gigerenzer and Hoffrage (1995) proposed two explanations to account for the facilitating effect of natural frequencies: computational simplification and evolutionary preparedness for (natural) frequencies. With probabilities, three pieces of information have to be taken into account, whereas with natural frequencies there are only two. Probability representations require the base rates (e.g., of disease and no disease) to be used to multiply the two likelihoods (e.g., probability of symptom given disease, and symptom given no disease). With natural frequencies, in contrast, the base rates are already contained implicitly and thus do not have to enter the calculation explicitly (Gigerenzer & Hoffrage, 1995). Further studies, however, showed that computational simplification alone cannot account for the increased performance of people using natural frequencies[8]. Instead, these authors conclude that reasoning performance increases substantially when information is presented in terms of the natural frequencies that correspond to the way organisms have acquired information through much of evolutionary history—that is, by naturally sampling (and tallying) events observed in the natural environment (see also Brase, 2002b).

Conclusions: The advantages of limited cognitive systems

Cognitive psychologists have long studied the limitations of human thought, and with good reason. Despite Hamlet's exhortation that we humans are "noble in reason...infinite in faculty" (Act 2, Scene 2), we struggle to keep more than a half dozen things in mind at once, we quickly forget what we have learned, we ignore much of the available information when making decisions, and we find it difficult to process deeply what information we do consider. But in focusing on the negative implications of these limitations, cognitive psychology may have grabbed the wrong end of the stick. The limited human mind is not just the compromised result of running up against constraints that can little be budged, such as the current birth-canal-limited size of the skull; rather, it is a carefully orchestrated set of systems in which limits can actually be beneficial enablers of functions, not merely constraints (Cosmides & Tooby, 1987). A less limited mind might fare worse in dealing with the adaptive problems posed by the structured environment. As Guildenstern later responded to Hamlet, presciently summing up modern psychology's computationally-intensive theories of cognition, "there has been much throwing about of brains." In many cases, throwing fewer brains at a task might do the trick—more is not always better.

---

[8] Consistent with this conclusion, Brase (2002a) has shown that frequencies were seen as clearer and easier to understand than single-event probabilities.

Considering the widespread selective pressures and attendant costs and benefits that have acted over the course of evolution on our cognitive mechanisms can help us to uncover these surprising instances when limitations are beneficial (and helps us understand the design and functioning of those mechanisms even when their limits are constraining). As we have seen in this chapter (see Hertwig & Todd, 2003, for more), limited information use can lead simple heuristics to make more robust generalizations in new environments. Forgetting in long-term memory can improve the performance of recall, and can protect individuals from harmful reactions at vulnerable periods in their lives. And limited short-term memory can amplify the presence of important correlations in the world.

But beyond just enhancing the abilities of certain cognitive systems, limits can even enable functions that may not be possible otherwise. According to Newport's (1990) "less-is-more" hypothesis on the role of limits in language learning, "the very limitations of the young child's information processing abilities provide the basis on which successful language acquisition occurs" (p. 23). Elman (1993) tested this idea with a neural network model, which he found was unable to learn the grammatical relationships in a set of several thousand sentences when given a large memory, but which could pick up the grammar incrementally if memory started small and gradually expanded. As he explained, "The initial memory limitations ... act as a filter on the input, and focus learning on just that subset of facts which lay the foundation for future success" (Elman, 1993, pp. 84-85).

These potential benefits of cognitive limitations compose one of the main themes we have sketched here in our picture of the issues that should be addressed within an evolution-inspired cognitive psychology. We have portrayed the importance of considering how general selective pressures—those arising in multiple task domains—can shape adaptive cognitive mechanisms, in addition to the shaping forces of domain-specific task requirements and environment structure. But most of the picture remains to be filled in. Here are some of the pressing questions open for immediate exploration (see Todd, Gigerenzer, & the ABC Research Group, 2000, for others): How is the adaptive toolbox of cognitive mechanisms filled—that is, what are the processes through which heuristics evolve, develop, are learned individually, or are acquired from one's culture? How do we select particular tools in particular situations? What role do other possible factors, such as emotions or social norms, play in decision heuristics? How effective can small information samples be for learning about one's environment? How does the use of particular cognitive mechanisms actually shape the environment itself (e.g., Todd & Kirby, 2001)? What selective pressures have shaped other cognitive capacities we have not touched upon such as attention, categorization, and planning?[9] And what methods are most appropriate for studying the action of selective forces on cognitive adaptations?

[9] For instance, the history of research on categorization reflects the rise of openness to evolutionary thinking in psychology, as it progressed from the use in the 1950s of artificial stimuli assigned to artificial categories by logical rules, to demonstrations in the 1970s that such research did not generalize to natural categories like species (Rosch, 1975), to arguments in the 1990s that human categorization is driven by domain-specific principles (Tooby & Cosmides, 1992; Hirschfeld & Gelman, 1994).

Taking an evolutionary perspective can help introduce new unity and coherence (as well as new ideas and hypotheses) into cognitive psychology. But the benefits of bringing the cognitive and evolutionary approaches to psychology together do not flow solely from the latter to the former. Cognitive psychology is also a salutary approach for evolutionary psychologists to engage with: It points to the importance of information, hence of the environment that it reflects, and the structure of the environment must be a central aspect of any evolutionary explanation of behavior. The field's experimental methodology is an important component of supporting and revising evolutionarily-inspired hypotheses regarding human cognition and action. Finally, cognitive psychology also reminds us of the crucial role that processing information with specific algorithmic mechanisms plays in the generation of adaptive behavior. This step—cognition—is often the "missing link" in non-psychological approaches to investigating the evolution of behavior (Cosmides & Tooby, 1987), and is still too often missing within evolutionary psychology studies, as in those that merely assert correlations between environmental cues and behavioral outcomes. By cross-fertilizing these two traditions, evolutionary and cognitive, a more vigorous hybrid psychology will be formed.

Acknowledgments

We are grateful to Gary Brase and Lael Schooler for helpful comments on this chapter.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J.R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703-719.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*, 396-408.

Anderson, J. R., & Schooler, L. J. (2000). The adaptive nature of memory. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 557-570). New York: Oxford University Press.

Ariely, D., and Levav, J. (2000). Sequential choice in group settings: Taking the road less traveled and less enjoyed, *Journal of Consumer Research*, *27* (3), 279-290.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417-423.

Barron, G., & Erev, I. (2002). Feedback-based decisions and their limited correspondence to description-based decisions. Manuscript submitted for publication.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognitions*, *5*, 176-196.

Brase, G. L. (2002a). Ecolgical and evolutionary validitiy: Comments on Johnson-Laird, Legrenzi, Girotto, Legrenzi, and Caverni's (1999) mental-model theory of extensional reasoning. *Psychological Review*, *109*,722-728.

Brase, G. L. (2002b). Which statistical format facilitates what decisions? The perception and influence of different statistical information formats. *Journal of Behavorial Decision Making*, *15*, 381-401.

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. & 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 (2),* 277-293.

Bullock, S., and Todd, P.M. (1999). Made to measure: Ecological rationality in structured environments. *Minds and Machines*, *9*(4), 497-541.

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind* (2nd ed.). Boston: Pearson.

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *Quarterly Journal of Experimental Psychology*, *52A*(2), 273-302.

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences, 7*(1), 19-22.

Christensen-Szalanski, J. J. J., & Beach, L. R. (1982). Experience and the base-rate fallacy. *Organizational Behavior and Human Performance*, *29*, 270-278.

Cosmides, L., and Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupré (Ed.), *The latest on the best: Essays on evolution and optimization* (pp. 277-306). Cambridge, MA: MIT Press/Bradford Books.

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

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In G. Gigerenzer, P.M. Todd, & the ABC Research Group, *Simple heuristics that make us smart.* New York: Oxford University Press.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95-106.

DePrince, A. P., & Freyd, J. J. (2004). Forgetting trauma stimuli. *Psychological Science*, *15*, 488-492.

Dhami, M. K., and Wallsten, T. S. (in press). Interpersonal comparison of subjective probabilities. *Memory & Cognition.*

Dudey, T., and Todd, P.M. (2002). Making good decisions with minimal information: Simultaneous and sequential choice. *Journal of Bioeconomics, 3,* 195-215.

Dukas, R. (ed.) (1998). *Cognitive ecology: The evolutionary ecology of information processing and decision making*. Chicago: The University of Chicago.

Dukas, R. (1999). Costs of memory: Ideas and predictions. *Journal of Theoretical Biology*, *197*: 41-50.

Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology*. New York: Dover.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic and A. Tversky *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge University Press.

Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition, 48,* 71-99.

Fiedler, K., & Kareev, Y. (2004). *Does decision quality (always) increase with the size of information samples? Some vicissitudes in applying the law of large numbers*. Manuscript submitted for publication.

Fischhoff, B. (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Eds.) *Judgment under uncertainty: Heuristics and biases* (pp. 335-335). Cambridge, UK: Cambridge University Press.

Freyd, J.J. (1983). Shareability: the social psychology of epistemology. *Cognitive Science, 7*, 191-210.

Freyd, J.J. (1990). Natural selection or shareability? [commentary] *Behavioral and Brain Sciences, 13,* 732-734.

Freyd, J.J. (1996). Betrayal trauma. The logic of forgetting childhood abuse. Cambridge, MA: Harvard University Press.

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

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. & Krüger, L. (1989). *The empire of chance: How probability changed science and everyday life.* Cambridge, UK: Cambridge University Press.

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

Gigerenzer, G. & Goldstein, D.G. (1999). Betting on one good reason: The Take The Best heuristic. In G. Gigerenzer, P.M. Todd, & the ABC Research Group, *Simple heuristics that make us smart.* New York: Oxford University Press.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Gigerenzer, G., Todd, P.M., and the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Goldstein, D.G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer, P.M. Todd, & the ABC Research Group, *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.

Goldstein, W. M., & Weber, E. U. (1997). Content and discontent: Indications and implications of domain specificity in preferential decision making. In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on judgment and decision making.* Cambridge: Cambridge University Press.

Hertwig, R., Barron, G., Weber, E., & Erev, I. (in press). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*.

Hertwig, R., Erev, I., Barron, G., & Weber, E. (in press). Decisions from experience: Sampling and updating of information. In K. Fiedler & P. Juslin (Eds.), *Information sampling as a key to understand adaptive cognition*. Cambridge: Cambridge University Press.

Hertwig, R., Fanselow, C., & Hoffrage, U. (2003). Hindsight bias: How knowledge and heuristics affect our reconstruction of the past. *Memory*, *11*, 357-377.

Hertwig, R., & Gigerenzer, G. (1999). The "conjunction fallacy" revisited: How intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making, 12*, 275–305.

Hertwig, R., Hoffrage, U., & Martignon, L. (1999). Quick estimation: Letting the environment do some of the work. In G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple heuristics that make us smart* (pp. 209-234). New York: Oxford University Press.

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*. West Sussex: Wiley & Sons.

Hey, J. D. (1981). Are optimal search rules reasonable? And vice versa? (And does it matter anyway?) *Journal of Economic Behavior and Organization, 2,* 47-70.

Hey, J. D. (1982). Search for rules for search. *Journal of Economic Behavior and Organization, 3,* 65-81.

Hirschfeld, L. A., & Gelman, S. A. (Eds.). (1994). *Mapping the mind: Domain specificity in cognition and culture.* Cambridge: Cambridge University Press.

Hoffrage, U. & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, *73*, 538–540.

Hoffrage, U. & Gigerenzer, G. (2003). How to improve the diagnostic inferences of medical experts. In E. Kurz-Milcke & G. Gigerenzer (Eds.), *Experts in science and society* (pp. 249–268). New York: Kluwer Academic/Plenum Publishers.

Hoffrage, U., Hertwig, R., & Gigerenzer, G. (2000). Hindsight bias: A by-product of knowledge updating? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 566-581.

Hoffrage, U., Lindsey, S., Hertwig, R. & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290,* 2261–2262.

Hoffrage, U. & Pohl, R. F. (Eds.) (2003). Hindsight Bias. Special Issue in *Memory* (Vol. 11, Issue 4/5, pp. 329–504).

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Dover.

Juslin, P., & Olsson, H. (in press). Capacity limitations and the detection of correlations: A comment on Kareev (2000). *Psychological Review*.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*, 430-454.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263-291.

Kareev, Y. (1995a). Through a narrow window—working memory capacity and the detection of covariation. *Cognition*, *56*, 263-269.

Kareev, Y. (1995b). Positive bias in the perception of covariation. *Psychological Review*, *102*, 490-502.

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

Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, *126*, 278-287.

Kenrick, D.T., Sadalla, E.K., & Keefe, R.C. (1998). Evolutionary cognitive psychology: The missing heart of modern cognitive science. In C. Crawford & D. L. Krebs (Eds.), *Handbook of evolutionary psychology* (pp. 485-514). Hillsdale, NJ: Erlbaum.

Klein, S. B., Cosmides, L., Tooby, J., & Chance, S. (2002). Decisions and the evolution of memory: Multiple systems, multiple functions. *Psychological Review*, *2*, 306-329.

Kleiter, G. D. (1994). Natural sampling: Rationality without base rates. In G. H. Fischer & D. Laming (eds) *Contributions to mathematical psychology, psychometrics, and methodology* (pp. 375–388). New York: Springer.

Koriat A., Goldsmith M. 1996b. Memory metaphors and the real-life/laboratory controversy: correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences*, *19*, 167-228.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481-537.

Krueger, J. I., & Funder, D. C. (in press). Towards a balanced social psychology: Causes, consequences and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences*.

Landauer; T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science, 10*, 477-493.

Lindsey, S., Hertwig, R. & Gigerenzer, G. (2003). Communicating statistical DNA evidence. *Jurimetrics: The Journal of Law, Science, and Technology, 43*, 147–163.
Marr, D. (1982). *Vision*. San Francisco: Freeman.
Martignon, L. & Hoffrage, U. (2002). Fast, frugal and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29–71.
MacGregor, J. N. (1987). Short-term memory capacity: Limitation or optimization? *Psychological Review*, *94*, 107-108.
McNally, R. J., Clancy, S. A., & Schacter, D. L. (2001). Directed forgetting of trauma cues in adults reporting repressed or recovered memories of childhood sexual abuse. *Journal of Abnormal Psychology, 110*, 151-156.
Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
Miller, G.F., and Todd, P.M. (1998). Mate choice turns cognitive. *Trends in Cognitive Sciences, 2,* 190-198.
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.
Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
Pirolli, P. L., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*, 643-675.
Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science*, *253*, 980-986.
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.* New York: Oxford University Press.
Rieskamp, J., & Otto, P. (in preparation). *SSL: A theory of how people learn to select strategies*. Manuscript submitted for publication.
Rosch, E. H. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192-253.
Roth, A. E., & Sotomayor, M. A. O. (1990). *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge: Cambridge University Press.
Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston: Houghton Mifflin.
Schooler, L., & Hertwig, R. (2004). *Environment and forgetting as natural allies of heuristic inference*. Manuscript submitted for publication.
Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the 'secretary problem'. *Organizational Behavior and Human Decision Processes*, *69*(3), 221–236.
Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior*, *6*, 156-163.
Simão, J., and Todd, P.M. (2003). Emergent patterns of mate choice in human populations. *Artificial Life, 9,* 403-417.

Simon, H.A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69,* 99–118.
Simon, H.A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–19.
Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology*, *25*, 207-222.
Tanner, C. & Medin, D.L. (in press). Protected values: No omission bias and no framing effects. *Psychonomic Bulletin and Review*.
Todd, P.M. (2000). The ecological rationality of mechanisms evolved to make up minds. *American Behavioral Scientist*, *43*(6), 940-956.
Todd, P.M. (2001). Fast and frugal heuristics for environmentally bounded minds. In G. Gigerenzer and R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (Dahlem Workshop Report), pp. 51-70. Cambridge, MA: MIT Press.
Todd, P.M., Fiddick, L., and Krauss, S. (2000). Ecological rationality and its contents. *Thinking and Reasoning, 6*(4), 375-384.
Todd, P.M., Gigerenzer, G., and the ABC Research Group (2000). How can we open up the adaptive toolbox? (Reply to commentaries.) *Behavioral and Brain Sciences*, *23*(5), 767-780.
Todd, P.M., and Kirby, S. (2001). I like what I know: How recognition-based decisions can structure the environment. In J. Kelemen and P. Sosík (Eds.), *Advances in Artificial Life: 6th European Conference Proceedings (ECAL 2001)*, pp. 166-175. Berlin: Springer-Verlag.
Todd, P.M., and Miller, G.F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer, P.M. Todd, and the ABC Research Group, *Simple heuristics that make us smart* (pp. 287-308). New York: Oxford University Press.
Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind* (pp. 19-136). Oxford: Oxford University Press.
Trivers, R. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man: 1871–1971* (pp. 136–179). Chicago: Aldine.
Tulving, E. Craik, F. I. M. (eds.) (2000). *The Oxford handbook of memory*. New York: Oxford University Press.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *2*, 105-110. Reprinted in D. Kahneman, P. Slovic & A. Tversky (Eds.) (1982), *Judgement under uncertainty: Heuristics and biases* (pp. 23-31). Cambridge: Cambridge University Press.
Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* (4) 281-299.
Von Neumann, J. (1958). *The computer & the brain*. New Haven, CT: Yale University Press.
Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk-sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 111,* 430–445.
Wimsatt, W. C. (1987). False models as means to truer theories. In: M. Nitecki and A. Hoffman (eds.), *Neutral models in biology* (pp. 23-55). London: Oxford University

Press.
Wixted, J. T. (1990). Analyzing the empirical course of forgetting. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *16*, 927-935.
Wixted, J. T. & Ebbesen, E. B. (1991). On the form of forgetting. P*sychological Science*, *2*, 409-415.
Wixted, J. T. & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition*, *25*, 731-739.