11

Human Memory Search

Eddy J. Davelaar and Jeroen G. W. Raaijmakers

Abstract

The importance of understanding human memory search is hard to exaggerate: we build and live our lives based on what we remember. This chapter explores the characteristics of memory search, with special emphasis on the use of retrieval cues. We introduce the dependent measures that are obtained during memory search, such as accuracy and search time, and discuss how these have contributed to our understanding of human memory search. The three phases of memory search (initiation, progression, and termination) are discussed in relation to the strategies employed by the human retriever. Finally, the experimental paradigms used in the memory literature are compared to examples of animal foraging behavior to identify points of contact for developing a general cross-domain understanding of search processes.

Introduction

In the cognitive sciences, human memory holds a special place. The ancient Greeks debated the origins and phenomenology of memory well before psychology existed as a recognized discipline. Within psychology, memory has a checkered past: it has been strongly connected to consciousness, been actively ignored during the behaviorist era, and has subsequently been reinstated as a bona fide topic of investigation. Despite the long history of research in human memory, many questions still remain and others have become more refined based on scientific advancements. In this chapter, we provide an overview of the cognitive components of human memory search.

The importance of understanding human memory search is hard to exaggerate. In everyday lives, people talk with each other about past events. During such conversations, information needs to be retrieved as quickly as possible and preferably be an accurate description of those events. The accuracy of retrieved memories is a critical aspect in legal court cases where prosecution of the defendant depends on eyewitness testimonies. These types of memories are referred to as episodes and episodic memory retrieval and will be addressed in this chapter. Retrieval from semantic memory (i.e., memory for facts and

From “Cognitive Search: Evolution, Algorithms, and the Brain,” edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

encyclopedic knowledge), such as conducted by a medical doctor when making a diagnosis based on observed medical test results, will also be addressed. These test results will trigger a search through semantic memory for their possible causes. Understanding semantic memory retrieval will lead to an understanding of how medical doctors derive with their list of candidate diagnoses. The speed and accuracy of episodic and semantic memory retrieval has secondary effects on processes that depend on them with potential life-altering or life-threatening consequences.

In the research laboratory, human memory is investigated using a range of memory tasks in a variety of paradigms. These tasks can be roughly categorized as single- versus multi-item recall, free-ordered versus serially ordered recall, and recognition. Episodic recall tasks involve reporting a single or multiple item(s) from a recently experienced event, such as words on a list. Semantic recall tasks involve reporting information from the long-term knowledge base, such as the exemplars of the category *animal*. The type of task used constrains the types of dependent measures that can be obtained. Common dependent measures used in research are memory accuracy, various indicators of retrieval time, confidence, and various derived measures related to memory organization. Here, we focus mainly on free recall paradigms, where information needs to be retrieved from episodic memory. This choice for episodic free recall is primarily due to its nature of allowing the participant maximal freedom to deploy search strategies. Where relevant, we contrast the search strategies with the more constrained semantic recall tasks.

We eschew the discussions on short-term or working memory and the comparisons of competing memory theories. Instead, we highlight the common views on memory search and point to productive areas for further research. We discuss characteristics of memory search and retrieval cues, which are hints that help memory retrieval, and focus on dependent measures that are obtained during memory search, such as accuracy and search time. The termination of an open-ended memory search constitutes our focus in the penultimate section, after which we address other approaches to human memory search that take inspiration especially from the animal foraging literature.

General Characteristics of Human Memory Search

To make our discussion of the general characteristics of memory search more concrete, we offer the following example: Suppose you are in a conversation and someone (let us call him Bob) mentions the film *Enemy of the State*. Bob tries to recall the name of the leading actor but is unable to do so. Let us now assume that you have not seen that particular movie. Bob mentions that the leading role is played by a black actor who is quite famous. You suggest "Denzel Washington" but Bob says no. After a while, you think of the city of Washington, D.C. and other U.S. cities, including Los Angeles. You suddenly

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

remember a TV series about a black family in Bel Air, which prompts you to say: "Oh, it must be Will Smith."

We have all experienced such a situation many times. It represents a prototypical situation of memory search and also reveals some of the key characteristics of memory search processes. First, what is retrieved from memory at a given moment is determined by the cues that are available. Second, through the use of such cues we have some control over what is retrieved from memory. Third, retrieval of the information that we are seeking may be hampered by other similar information in memory (in this example, the name Denzel Washington). Finally, although not in our example, we may give up at some point and decide that further search is useless.

The above example represents a semantic-cued recall task, in which hints (black actor, famous, not Denzel Washington) are given, while you jog your memory. Cues are not always provided by the environment in this manner, and people may generate their own cues (Los Angeles). In contrast, the prototypical paradigm used in the research laboratory is the (episodic) list-learning paradigm, in which a series of items (pictures, words, letters) are presented one at a time to the participant for memorization. After presentation of this list, the participant is asked to report all items that were memorized: the larger the number of items on the list, the lower the probability of recalling each item (e.g., Murdock 1962). Interestingly, the order in which the items are reported and the retrieval latencies reveal much about the search strategies employed by the participant on this task. Generally, memory search is characterized by the use of retrieval cues, the three stages of memory retrieval, and by its sequential, self-terminating nature, which we now discuss in turn.

The Cue-Dependent Nature of Memory Search

Tulving and Madigan (1970) once characterized memory retrieval using the Latin proverb: *Ex nihilo nihil fit* (nothing comes from nothing). This saying points nicely to the critical importance of *retrieval cues*. It is a common assumption in human memory research that when something comes to mind, there is always a triggering stimulus. This may be an external event, such as a question that is asked or a specific remark, but it may also be an internally generated event, such as a particular thought. Over the past thirty years or so, memory researchers have especially emphasized the importance of *context* as a retrieval cue. Thus, being in the same environmental context as during the original event or being in the same physiological state helps the retrieval of information stored in that context. This has been called the *encoding-specificity principle*, which states that successful retrieval is a function of the overlap between the information present at retrieval and the information stored in memory (Tulving and Thomson 1973).

The importance of retrieval cues may be understood if one assumes that what gets stored in a memory trace is a sample of the information that was

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

present in the mind's eye at the time of the original event. Hence, the memory trace includes not just the target event but also any fleeting thoughts and feelings that happen to be present. It is generally assumed that the features present in the retrieval cues determine what is activated from memory. Hence, retrieval cues have two (related) properties: they determine which memory traces are activated (i.e., determine which traces are in the search set) and how strongly a trace within the search set is activated. It is often assumed that the higher the overlap in the features present in the retrieval cues and the stored trace, the more that trace will be activated (Tulving and Thomson 1973); in some recent models, activation is a function of both the number of overlapping and nonoverlapping features (see Shiffrin and Steyvers 1997).

Strategic and Automatic Aspects of Memory Retrieval

Generally, memory search consists of three phases: initiation, progression, and termination. Given a specific set of retrieval cues, the retrieval process is completely automatic in that the activation of memory traces, given a specific set of cues, is an automatic process that is determined by the associative strengths from the cues to the memory traces. This does not mean, however, that we have no control over what is retrieved from memory. Each of the three phases is under strategic control. For example, we have some strategic control over what is retrieved through the choice of retrieval cues. When we are trying to recall a specific name, we may resort to an alphabetic strategy, simply trying the successive letters of the alphabet to see whether one "works" (see Gronlund and Shiffrin 1986). In addition to the choice of retrieval cues, which affects the progression of memory search, there are additional aspects where there is some strategic control, two of which have been discussed in the literature (Raaijmakers and Shiffrin 1981). First, before the actual search process, there is the decision to search or not to search. We may decide on the basis of the information given that a memory search is unlikely to lead to a successful answer and decide not to even make an attempt. It is usually assumed that such a choice is based on a quick evaluation of the amount of activation generated by the available cue information. If this falls below some criterion, we may quickly decide that the answer is unlikely to be found so that a search process would be futile. Second, after an unsuccessful search attempt, we have a choice to either give up or continue the search. If we continue searching, we may decide to change the set of probe cues used (e.g., by including information retrieved on prior search attempts) or maintain the same set of cues.

Memory Search as a Sequential, Self-Terminating Process

Many models for recall are formulated in such a way that the probability of successful recall is given by some analytic formula (some function of, e.g., the study time, the retention interval, and/or the strength of competing memory

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

traces). However, analyses of retrieval latencies (discussed further below) reveal that recall memory is well captured by the assumption that recall is based on a sequential process (Diller et al. 2001; Nobel and Shiffrin 2001). Thus, in theoretical models of recall memory (Raaijmakers and Shiffrin 1981), it is assumed that the retrieval process consists of a series of retrieval attempts. Each retrieval attempt may end with either successful recall, a decision to stop (give up), or a decision to continue the search process. Importantly, the decision to terminate the search process is based on the unfolding of the search itself. In other words, there are typically no external criteria, such as a fixed time limit for retrieval, by which search is terminated. In unlimited time, memory search in recall tasks is self-terminating. We will come back to how terminating decisions are reached later.

Characteristics of Retrieval Cues

Types of Retrieval Cues

As mentioned earlier, retrieval cues consist not just of the test item as presented to the subject, but also of various other types of information. In our example of naming the leading actor in *Enemy of the State*, the additional retrieval cues can be the thoughts generated during the memory search (e.g., "Denzel Washington," "Los Angeles"). Within the list-learning paradigm, these other types of information may further include other items that were presented in close temporal proximity to the target item (e.g., another item on the list), items that have a preexperimental association to the target item (e.g., extralist cues), things the subject thought about while encoding the item (e.g., a mental image formed to connect the list items), the internal physiological state (e.g., if the subject was under the influence of a particular drug during encoding), and the external context (e.g., the room in which the encoding took place).

Each of these types of cues has been shown to affect the probability of retrieving the target item. In list free recall, recalled items are most likely to be followed by recalling other items from neighboring serial positions (Howard and Kahana 1999; Kahana 1996), indicating that one item can cue another nearby in the list. Performance in recall tasks is higher when the physiological state corresponds to the state the subject was in during encoding (Eich 1977, 1980). This has been shown for both emotional states as well as for drug-induced states, even when drugs by themselves have a negative effect on memory. For example, even though alcohol by itself has a negative effect on memory, recall performance is better after (moderate) consumption of alcohol if the encoding also took place while being under the influence of alcohol (Goodwin et al. 1969). Similarly, testing in the same environmental context has a positive effect on recall (Godden and Baddeley 1975). This even holds if the testing is

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

done in a different context but the subject is reminded of the encoding context (e.g., by giving a photograph of the original context; Smith 1979).

Effectiveness of Retrieval Cues

Whether or not a specific retrieval cue is effective depends on a number of factors. The two most important ones are (a) the strength with which the cue is associated to the target item, and (b) the number of other terms that are also associated with the cue.

The first factor corresponds to what is often termed "memory strength" and is considered to be a function of the number of matching features between the cue and the memory trace, and possibly the number of mismatching features (see the REM model; Shiffrin and Steyvers 1997). The second factor has been termed the size of the search set. The search set can be defined as the set of retrieval candidates that are activated in response to retrieval cues. An effective retrieval cue will be one that limits the search set to a few memory traces (including, of course, the target trace). The opposite happens when cues activate a large set of distracting traces (Anderson 1974; Watkins and Watkins 1976). There also appears to be a general rule such that if a cue does not lead to additional focusing of the search (decreasing the size of the search set) that cue will not lead to an increase in memory performance.

Cue Combination

Given specific values for the associative strengths of a cue to the memory traces, memory models specify some kind of rule to translate these strengths into a predicted probability of recall. This could be some negative exponential function (as in the ACT model; Anderson 1983) in which absolute strength is transformed into a probability of retrieval, or a relative strength calculation in which the absolute strength is divided by the total strength of all competing memory traces (e.g., Raaijmakers and Shiffrin 1981). A separate issue, however, concerns what the relation should be when two (or more) cues are combined. Should the search set consist of all traces associated to either of the cues, or only those traces associated to both cues? An empirical answer to this question was obtained by Humphreys et al. (1991). In a semantic memory task, they showed that when two cues were given in combination, the search set (the number of items compatible with both cues) was limited to exactly one and the probability of retrieving the answer was very high; when either cue was given by itself, retrieval probably was much lower. Nairne (2002) points out that to be effective, the additional cues should have more overlap with the target item than with distracting information.

From these findings we may conclude that the search set should be more or less equal to the intersection of the search sets evoked by each cue separately. Such a mechanism explains why the probability of finding an answer increases

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

as we are given more relevant (and diagnostic) information (more retrieval cues). Formal models of memory, such as the *search of associative memory* or SAM (Raaijmakers and Shiffrin 1980, 1981), have incorporated such a mechanism through the assumption that the activation of a trace, when multiple cues are given, is a function of the (weighted) product of the associative strengths to each cue separately.

Cue Switching/Updating

Although in some search models it has been assumed that the cues are determined at the start of the search process, a more realistic assumption is that information that is retrieved during the search process may subsequently be used as an additional cue or may replace one of the cues used thus far. In the SAM model for free recall, for example, it was assumed that an item recovered during the search would then be used as an additional cue to make it easier to retrieve items that had been associated to the recovered item during the initial study. Similarly, in the models proposed by Kahana and his associates (see Howard and Kahana 1999), it is assumed that contextual information retrieved during the search is used to update the contextual retrieval cue used in subsequent searches. In many models of serial recall, contextual cues represent the list position and are updated during the retrieval process (Brown et al. 2000; Burgess and Hitch 1999; Henson 1998). Therefore, these models assume that the unfolding of context information during retrieval is independent of the retrieved items. Contemporary models that tackle both serial and free recall tasks (Anderson et al. 1998; Brown et al. 2007a) have yet to resolve this distinction.

The updating of cues based on retrieved information leads to clustering across recalled items. For example, in episodic free recall tasks, in which list words are drawn from multiple semantic categories, participants tend to use the category label (which they generate internally as recall proceeds) as a retrieval cue to recall list items in clusters, first from one category, then another, and so on (Patterson et al. 1971; Tulving and Pearlstone 1966). This pattern is also seen in the semantic fluency task, in which participants are asked to report as many animal names as possible in a given time (Bousfield et al. 1954; Bousfield and Sedgewick 1944). Participants tend to cluster the animals by subcategories, such as zoo animals, pets, and aquatic animals. Thus, cue updating/switching seems to be present in searching for multiple items in both episodic and semantic memory. This search behavior is what Hills and Dukas (this volume) refer to as area-restricted search in internal environments (see also Stephens et al., this volume).

The three phases of memory search (initiation, progression, and termination) apply to each retrieval cue, making the entire memory search consist of a hierarchy of cue-related retrieval. Both levels of the hierarchy have a signature in memory accuracy and retrieval time to which we turn next.

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

Memory Accuracy

From the beginning of experimental investigation into human memory search, the focus has predominantly been on the accuracy (and its derived measures) of memory search, possibly due to the easier method of collecting such data for accuracy than for latencies. Despite the wealth of data of memory accuracy, questions and debates exist regarding the processes that underlie the data. Here we highlight three such findings for the list-learning paradigms.

First, when participants are instructed to report any items that come to mind, in addition to retrieving items from the list, participants may (a) retrieve items from previous lists, (b) report items related semantically or phonologically to target items, and (c) repeat items that were already reported. With normal instructions, these errors are relatively rare. Evidence for this view comes from studies (Kahana et al. 2005; Unsworth et al. 2010) in which participants reported anything that comes to mind during the retrieval period. In those studies, large numbers of intrusions are produced that are related to the items from the list. Thus, without the explicit instruction to report everything that comes to mind, a filtering process occurs after generation of the items. Despite its importance, this filtering process has yet to be unraveled. Initial attempts involve using existing processes, such as recovery in the SAM model, as the locus for memory filtering (Kimball et al. 2007; Sirotin et al. 2005).

Second, memory accuracy can be conditionalized as a function of the input list position, leading to serial position profiles with increased accuracy of free recall for items from the beginning and end of the list (primacy and recency effects, respectively). The common view is that primacy effects are due to extra rehearsal of the early items (but see Tan and Ward 2000). A long-standing debate questions whether recency effects in immediate free recall reflects retrieval from a short-term store. Formal models that argue against the existence of a short-term store (Brown et al. 2007a; Howard and Kahana 2002) attribute all recency effects to the encoding-retrieval match combined with changing episodic context. Because the context gradually changes during encoding of the list items, recent items are encoded in a context that is more similar to the context at retrieval than are earlier items, and this gradient of contextual similarity underlies the recency effect. Models that also include a short-term store (Atkinson and Shiffrin 1968; Davelaar et al. 2005; Raaijmakers and Shiffrin 1980, 1981) attribute recency effects in immediate free recall to a more accurate readout from the short-term store. The debate centers around the need to postulate a short-term store to account for data (for reviews, see Davelaar et al. 2005; Sederberg et al. 2008).

Third, memory accuracy can be conditionalized against the distance between the serial positions of the previously retrieved item and the current item. A robust finding in free recall is the observation that successively reported items were presented in close proximity during encoding (Kahana 1996). In other words, when retrieving a word from list position *n*, the next word that

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

is retrieved is more likely to be from position $n + 1$ or $n - 1$ than $n + 2$ or $n - 2$. This supports models that include a changing context representation (Estes 1955; Howard and Kahana 2002; Mensink and Raaijmakers 1988). In addition, this so-called lag-recency effect is asymmetric with the forward transitions (e.g., $n + 1$) being more likely than backward transitions (e.g., $n - 1$). This asymmetry has been explained in terms of preexperimental context being retrieved and incorporated in the ongoing changing context during encoding and retrieval (Howard and Kahana 2002). Detailed predictions from these assumptions are still heavily debated using formal modeling (Farrell and Lewandowsky 2008).

Memory Retrieval Time

Memory search takes time, and the profile of memory search latencies have been used to address a number of questions regarding the dynamics of retrieval. Here, we review some temporal variables and their impact on theorizing.

Cumulative Retrieval Functions

As early as the 1940s, researchers focused on the cumulative recall function (Bousfield and Sedgewick 1944). This function sets the total number of items retrieved so far against the time spent in memory search. This function is shown to be well described by a cumulative exponential:

$$N(t) = N_{asy}{}^{*}\left[1 - \exp(t/\tau)\right], \tag{11.1}$$

with recall asymptote N_{asy} and mean latency τ (Bousfield and Sedgewick 1944). Researchers theorize that this good fit of a cumulative exponential is indicative of a system in which items are sampled with replacement, tested, and reported if they have not already been retrieved (Indow and Togano 1970). If there is a finite-sized pool of retrieval candidates, sampling-with-replacement leads to a diminishing rate of sampling-yet-unretrieved items.

The sampling-with-replacement process has become a critical element in theories of human memory search. The important assumption is that to obtain an exponential cumulative retrieval function, retrieved items should be independent of each other, which is not the case when, for example, retrieved items are clustered. Deviations from exponential functions have been observed and discussed in terms of the dependence among retrieved items, both in semantic retrieval tasks (Bousfield et al. 1954) and episodic retrieval tasks (Patterson et al. 1971). In addition, deviations from exponential functions are also observed when participants employ specific strategies within cued categories (Gronlund and Shiffrin 1986; Indow and Togano 1970).

From “Cognitive Search: Evolution, Algorithms, and the Brain,” edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

The actual process by which resampling can occur is still unclear. Three options can be discerned in the literature. First, the sampled item remains activated to the full extent given the retrieval cue (Davelaar 2007; Indow and Togano 1970). Second, the sampled item receives a decreased sampling probability that is still above some baseline level. This approach features in models of serial recall, which use a competitive queuing process (for a review and comparison, see Davelaar 2007) that is employed to produce sequential output. Third, the sampled item is increased in strength, making it more likely to be resampled. This increment is explicitly modeled in SAM, as a free parameter, allowing SAM to hover between the first and third options. Related to the third option is the proposal that a sampled item is re-encoded in memory, but in a separate trace (Laming 2009; Nadel and Moscovitch 1997). As retrieval continues, this sampled item has an increased opportunity to be resampled, even though the strength of each trace is unaltered. The different options do make different predictions with regard to retrieval latencies, which future research may elucidate.

Interresponse Times

In addition to a global cumulative retrieval function, the time between successive retrievals is a further temporal variable of great importance in studies of memory search. Several studies have analyzed the intricacies of interresponse times (IRTs) (Murdock and Okada 1970; Patterson et al. 1971; Rohrer and Wixted 1994; Wixted and Rohrer 1993, 1994). The main finding is that in episodic retrieval, the IRTs increase with more items retrieved. Rohrer and Wixted (1994; see also Rohrer 1996) presented evidence to suggest that the IRTs follow a pure-death hyperbola, in which the mean ith IRT equals the mean retrieval latency τ (across all items) divided by a number of items still in the finite-sized search set. This inevitably implies that at any given time the size of the search set can be estimated by the size of the IRT.

The validity of estimating search set size from IRTs was initially tested by manipulating list length, presentation duration, and proactive interference (Rohrer 1996; Rohrer and Wixted 1994; Wixted and Rohrer 1993). The method was subsequently applied to verify the loss of memory traces in patients with Alzheimer's disease (Rohrer et al. 1995) and the decreased rate of retrieval in patients with Huntington's disease (Rohrer et al. 1999). IRTs have been found to be sensitive to whether items are retrieved from episodic or semantic memory (Rohrer 2002), suggesting that the relations among items need to be considered in deriving conclusions based on IRTs. This is most prominently demonstrated in the categorized recall task (Patterson et al. 1971) mentioned earlier, in which within-cluster IRTs are much faster than between-cluster IRTs.

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

Total Time and Exit Latency

Apart from the temporal microdynamics, two further measures of retrieval time have been utilized in recent years. Dougherty and Harbison (2007) modified the standard free recall paradigm to allow participants to indicate when they have finished memory search. The instructions were given before the experiment, allowing participants to calibrate their internal system for the task. This slight modification produces a measure of total search time (i.e., the time from the start of the recall phase to the time of stopping) and the exit latency (i.e., the time between the onset of the last retrieved item and the time of stopping). The total time increases while the exit latency decreases with the number of items retrieved. These two additional measures have proven to be vital in our understanding of how memory search terminates, as we describe next.

Memory Search Termination

After a series of retrieval events, a person may decide to terminate memory search. In list recall tasks in the laboratory, participants may have various reasons to stop searching memory, such as wanting to receive the experiment payment for minimal effort, a lack of desire to help out in research, or a genuine feeling that further memory search will not lead to retrieving any more list items. When decisions, such as making a medical diagnosis, depend on shortlisting potential candidates, prematurely terminating the memory search for those candidates may have dire consequences (but so may searching too long in time-critical cases). Finding out how a person decides that further memory retrieval is futile relies on new paradigms and analyses, and may involve incorporating ideas from related areas, such as decision making.

Stopping Rules in Memory Search and Decision Making

In research on decision making, stopping rules are seen as an important factor in deciding effectively. Essentially, a stopping rule is needed to terminate an ongoing process (e.g., searching for information) so that a response can be generated. Here lies the important difference between the vast literature on stopping rules in decision making and the limited literature on stopping rules in memory search.

Browne and Pitts (2004) make a distinction between choice problems and design problems. Choice problems are characterized by the goal of choosing one out of several candidates. When the process stops, a single response is registered. Design problems are characterized by the goal of producing as many new responses as possible. When the process stops, the retriever has decided that further search will not produce any new responses. Problems studied in the decision-making literature are often of the first type, although there is

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

considerable research on decisions made in what could be considered a hybrid fashion, first involving searching for one or more cues, with stopping rules indicating when enough cues have been found to select a single option (Gigerenzer et al. 2012). Memory paradigms such as recognition memory, which require a single yes/no response, are also categorized as choice problems. This is in contrast with the recall paradigms discussed above which involve the generation of multiple responses and can thus be categorized as design problems.

Sophisticated methods exist to investigate the type of stopping rules used in memory tasks that require a single response, such as the *systems factorial technology* (Fific et al. 2008; Townsend and Nozawa 1995; Townsend and Wenger 2004). These methods have yet to be further developed to deal with stopping rules in memory paradigms that require multiple responses. In those tasks, participants employ a stopping rule aimed at producing as many items as possible. When memory search is terminated, the retrieved information may be used in a second step that involves selection among the retrieved items. Thus in a medical decision-making task, the medical doctor will employ two separate stopping rules: one for memory search to maximize the number of candidate diagnoses retrieved (a design problem) and one for selection to maximize the accuracy of final diagnosis (a choice problem).

Optimal Stopping

To understand the problem faced by the human retriever, it is useful to set memory search for multiple items against a wider set of related stopping problems that may inform optimal rules for design problems. The first is the rank-based sequential decision-making task, commonly known as the "secretary problem." In this task, a person interviews and ranks secretaries, one at a time. After each interview the person has to decide whether to hire the just-interviewed secretary or continue to the next one. Once a decision to continue is made, this secretary is taken off the list of candidates. The stopping rule aims to maximize the probability of hiring the best secretary out of those interviewed. As stopping results in a decision, the problem is a choice problem (for a review, see Freeman 1983). Even though people might use a satisficing (aspiration-level-based) stopping rule for both memory search and rank-based sequential decision making, the structural similarity between the problems is low. The best secretary may be anywhere in the sequence of interviewees, whereas the memory item that best matches the cue(s) will be activated most strongly and thus retrieved first.

The second stopping problem comes up in the capture-recapture approach to estimating the size of a population. Here, an animal is captured from a finite-sized population, marked and returned to the population (sampling-with-replacement), and the probability of recapturing the marked animal can be used to estimate the population size. The most useful rule for deciding when to stop

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

capturing the animals weighs minimizing the cost of capturing animals against the benefit of having a better estimation (for a review, see Nichols 1992). A related capture problem occurs in a debugging procedure in computer science (e.g., Chao et al. 1993; Forman and Singpurwalla 1977). Though such capture problems are similar to memory search in focusing on the yield of found versus unfound items, the details and aims of these problems make them choice problems (a decision about number of animals or remaining bugs is made) rather than design problems. Furthermore, in the bug capture problem, the important difference from memory search is the low probability of occurrence of bugs and the need to take the bug out of the pool of program code (sampling-without-replacement). Nevertheless, the requirement of capture problems to estimate the number of yet-to-be-captured targets may also be important in memory search, though new studies are needed to determine whether people actually make such estimates when retrieving items from memory.

These related stopping problems serve to emphasize the importance of explicitly defining the problem that humans face when retrieving information from memory. The assumptions drawn from the memory literature are that the retrieval process in list-learning paradigms and semantic fluency tasks involves sampling-with-replacement and the aim is to maximize the number of items retrieved while minimizing costs, both violated by the stopping problems just presented. On the other hand, the problems of rank-based sequential decision making and capture-recapture assume that individual candidates are independent of other candidates. For memory search this is an untenable assumption, given the episodic contextual association in all episodic recall tasks and semantic associations in all semantic recall tasks. The influence of the associative structure on stopping rules is a topic for future investigation.

Evaluating Stopping Rules in Human Memory Search

Four stopping rules commonly used in models of free recall were addressed by Harbison et al. (2009). These rules involved thresholds on:

1. total time spent retrieving (Davelaar et al. 2005);
2. time since last retrieved item (Rundus 1973);
3. decrease in retrieval rate (Young 2004); and
4. number of retrieval failures (Raaijmakers and Shiffrin 1980).

To test these stopping rules, Harbison et al. (2009) implemented these rules in the SAM memory model framework and quantitatively fitted the resulting models to data on total retrieval time and exit latency obtained from an open-ended free recall paradigm. The first three rules did not provide qualitative fits to the data. The number-of-failures rule captured the data qualitatively and also provided a strong quantitative fit.

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

The computational work by Harbison et al. (2009) showed that many computational theories use an implausible stopping rule for free recall. This is not to say that Rule 4 is the true stopping rule. Instead, in the absence of alternative rules that provide such quantitative fits, the number-of-failures rule is the best rule we currently have to describe how humans terminate their memory search. This can be compared with similar evaluative studies (e.g., Wilke et al. 2009) of stopping rules for cognitive search (as opposed to list recall). The deployment of Rule 4 in large-scale models of decision making, such as HyGene (Thomas et al. 2008), also gives better fits to human data on medical decision-making tasks.

An interesting observation is that toward the end of a recall protocol, participants tend to repeat already-reported items (Unsworth et al. 2010). Although this finding is striking, some (Laming 2009) hold that the participant's realization that the same word has already been retrieved triggers the decision to terminate memory search. Further empirical and computational work is needed to address the true causal relationships underlying increased repetitions and search termination.

Other Approaches to Human Memory Search

Our discussion in the preceding sections focused mainly on the mechanisms involved in memory search. These approaches use detailed analyses of memory accuracy and retrieval times. Alternative approaches provide powerful metaphors and analytic tools to further research in mechanisms of human memory search.

Keyword-Based Search Analogy

Human memory search is often likened to how information is retrieved from a database using a search engine with search terms combined by Boolean logic (e.g., *AND* and *OR*). There are, of course, many technical differences regarding the storage and retrieval of information; more informative differences between human memory search and keyword-based search are in terms of the use of cues and keywords. Typically in a search engine, typing keywords *A AND B* will produce information that is associated with both *A* and *B*. Humans, however, may still report *A-notB* items and *B-notA* items. Whereas these intrusions may seem to reveal limitations of the human memory search process, they crucially highlight the utilization of cues. For example, humans seem to interpret *A AND B* as *A OR B*, with a greater weight for *A-and-B* items. In sampling models (e.g., Raaijmakers and Shiffrin 1981), cues are combined multiplicatively, but it is not inconceivable to use an additive rule in which cues are differentially weighted. This would allow modeling the intrusions seen in humans together with the ability to select items that are associated with both cues.

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

Rational Analysis of Memory Search

A rational approach to human memory can be applied in ways that are similar to what has been done in the decision-making literature. One prominent example is the work by John R. Anderson and colleagues (Anderson and Milson 1989; Anderson and Schooler 1991). In their work, the retrieval of a memory trace is governed by two main factors: a history factor, which describes the pattern of prior use of the memory trace, and a context factor, which underlies the cue-dependency of memory retrieval. These two factors are multiplied to obtain the odds that the particular item is needed and thus will be retrieved. Human memory search is assumed to terminate when the need odds fall below a cost-benefit ratio. Thus far, the theory has been applied to the macrodynamics of memory retrieval, but a full rational analysis that includes temporal microdynamics is yet to be developed.

Animal Foraging

In recent years, researchers have compared search through the cognitive system with animal foraging behavior (Hills 2006; Hills et al. 2009; Hutchinson et al. 2008; Metcalfe and Jacobs 2010; Wilke et al. 2009). This is a very useful comparison and has allowed the wide literature on optimal animal foraging behavior to be integrated with cognitive search. To appreciate the similarities and differences, we recast two memory paradigms in terms of an animal foraging paradigm. The reader is invited to compare these examples with the chapters in this volume by Stephens et al. and by McNamara and Fawcett.

Cued recall memory paradigms could either involve one or multiple cues that result in one or more target items being retrieved. In human memory search, a cue demarcates the search set from which items are retrieved. Therefore, the search set can be compared to a patch of food, with the food items being analogous to the memory items. A task such as semantic fluency (naming as many animal names as possible) can be recast as foraging in a patch of *pets*, then a patch of *zoo animals*, then a patch of *aquatic animals*, and so on (see Hills et al. 2009). In our movie example presented earlier, the cues (black actor, famous) initially pointed to a wrong target item (Denzel Washington). This incorrect item changed the cognitive landscape (via U.S. cities) and opened up a path to the correct patch (a TV series set in Los Angeles) which involves the target item (Will Smith). Given such examples, one can address the question of whether patch-leaving behavior of animals is similar to cue-switching behavior in memory search. Hills et al. (2009) did exactly this and successfully applied a model of patch-leaving behavior to search through semantic memory.

Episodic recall tasks in lab settings typically require an initial step of learning a sequence of patches. These patches may contain a single item or a number of items. The size of the patches is determined during encoding, where strategies, such as rehearsal, lead to larger patches. These patches are

From “Cognitive Search: Evolution, Algorithms, and the Brain,” edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

connected by episodic links that are expected to follow a contextual similarity gradient. Assuming that all items are not semantically related, the paths among all learned items are only of an episodic nature. At retrieval, a searcher "forages" for the list items using the episodic paths. In free recall, the searcher forages through patches in any order; however, in serial recall, the searcher essentially exhibits trail following along the similarity gradient. Animals that use trail following, such as ants, leave behind chemical traces that gradually fade with time. The memory literature shows that the longer the list of items, the less likely the participant reports the items in serial order, and instead starts retrieval with more recent items (Ward et al. 2010). Therefore, to make the trail-following analogy of episodic recall work, one can hypothesize that episodic traces fade with time (though this raises the question of what to do if the trail fades away completely). However, this is inconsistent with work showing that the contextual gradients remain for a very long time (Howard and Kahana 1999). This example highlights limits to how widely ideas from the animal foraging literature can be applied to human memory search.

Apart from these aspects of the traces that link patches, an obvious distinction between animal foraging and human memory search is that in the latter, all the items have been experienced at least once, whereas animals may search for never-experienced food patches. Therefore, human memory search may be better compared to exploitation behavior in animals. Of course, not all human memory phenomena will be usefully comparable with animal foraging. For example, recognition memory involves a single yes/no response to a probe based on an overall sense of familiarity. This does not appear to involve a search process (Diller et al. 2001; Nobel and Shiffrin 2001) and thus cannot be reasonably compared with animal search behavior.

Information Foraging Approach

An approach to searching for information that is directly inspired by the animal foraging literature and attracts wide attention is information foraging theory (Pirolli 2007; Pirolli and Card 1999). According to this theory, a forager enters a patch of information and stays within that patch until the benefit of staying within that patch (in terms of the rate of gain of valuable information per unit time) falls below the benefit of searching elsewhere. The information foraging approach can be applied to memory search by assuming that each patch represents a subsearch set that is delineated by a retrieval cue. Recent work (Hills et al. 2009; Rhodes and Turvey 2007) suggests that this approach is useful in accounting for the clustering behavior seen in semantic memory retrieval that is known to defy the strict cumulative exponential retrieval function (Bousfield et al. 1954). Applying information foraging theory to more fine-grained temporal dynamics is one of the challenges for the near future.

From "Cognitive Search: Evolution, Algorithms, and the Brain," edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

Concluding Remarks

Over the last 100 years we have seen a remarkable increase in our understanding of how humans search for and retrieve information from memory. We are able to infer, based on profiles of memory accuracy and retrieval times, how the information is organized in the cognitive system and how it is found again. This increased understanding has helped in applications ranging from verifying claims about memory structure in patients with brain damage to shedding light on what makes individuals differ in their memory abilities. It has also fueled healthy debates on the precise interpretations of findings, which in turn has led to a deeper insight in the boundary conditions of particular theories. It is fair to say that without inspiration from considering the diversity of search strategies seen in humans and animals, the study of human memory search would have settled on a single cumulative exponential function of retrieval. The cognition is in the details.

From “Cognitive Search: Evolution, Algorithms, and the Brain,” edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.

From “Cognitive Search: Evolution, Algorithms, and the Brain,” edited by Peter M. Todd, Thomas T. Hills, and Trevor W. Robbins. 2012. Strüngmann Forum Report, vol. 9, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01809-8.