Optimal Foraging Theory: Foraging in the Face of Uncertainty

[Modeling, Psychology, Philosophy]

Abstract

Classic Optimal Foraging Theory (OFT) originally outlined food finding behavior through the Marginal Value Theorem (MVT), which describes the best time to move onto a new resource patch. The extension of the MVT to cognitive processes has allowed for the testing beyond the ideal observed behavior and into mechanistic processes that give rise to foraging behavior. At the core of the theory is the definition of the Optimal Forager, which place the assumptions necessary to describe the generalizability of the MVT. The original Optimal Forager definition held an orientation towards a behaviorist epistemology which allowed for high generalizability of the behavior. On the other hand, testing foraging as a cognitive process allowed researchers to ask questions beyond the ideal Optimal Forager definition and focus on variability in habitats as well as uncertainties foragers face in relevance to internal mechanistic processes. In this paper we will review canonical OFT and how it has been expanded into cognitive processes. We will also break down the assumptions that make the model work and in which cases those assumptions are limited. An argument is presented for defining a Heuristic Forager, which is used to formally update foraging epistemology trends in cognitive oriented research that was not set in the original Optimal Forager definition.

Intro

The problem of search shows itself across disciplines and varies widely in model implementations. Researchers have pointed out that search processes are behaviors that can generalize across species and types of habitats. On one hand, search behavior can be seen in plant roots trying to localize water, bacteria seeking nutrients, birds hunting for fish in the ocean, or early humans searching for edible berries in bushes (Koshland 1980; Weimerskirch, Salamolard, Sarrazin, & Jouventin, 1993; Eisenbach & Lengeler 2004; Raichlen et al., 2014). On the other hand, search behavior can also be seen when looking for a mate, searching your memory for a specific word, looking for hotel prices online, or when babies look for parental feedback through their babbling (Hills, Jones, & Todd, 2012; P. Pirolli & Card, 1999; Prokopy & Roitberg, 1984; Tamis-LeMonda, Kuchirko, & Song, 2014). The former presents classic cases in the behavioral ecology literature that can be framed as food finding behavior, where the goal of the search is to replenish energy to the searcher. The latter provides examples in which instead of energy consumption, the currency of trade in the search process is time. The principles of search behavior in the classic examples have been able to generalize beyond food finding behavior despite the differences in habitat, currency, and goal. This has been seen across varying types of search processes, but this paper will focus on the act of foraging. Broadly speaking, foraging involves a searcher exploring a space in order to exploit resources. When considering a balance of currency, whether it is time or energy, resources are modeled to be clustered in resource patches. This is referred to as patch foraging (Hills & Dukas, 2012).

The problem of a forager in patch foraging is to determine the best way to allocate its currency as it travels from patch to patch. The two state (exploration & exploitation) switch cost

balancing problem falls under the category of Optimal Control Problems. Problem sets in this category are assumed to have guidelines in which a best (optimal) solution is attainable (Kirk, 2012). In this paper we review the canonical theoretical framework for patch foraging and its model implantation for describing it as a type of Optimal Control Problem. We then demonstrate limitations to its categorization as an Optimal Control Problem by breaking down the assumptions necessary for the original model implementation. This paper concludes by providing an updated way of talking about the extension of foraging in cognitive processes.

OFT

The traditional Optimal Foraging Theory (OFT) was initially grounded in animal food finding behavior. It framed a scenario in which the food was found in patches of different magnitudes, and the predator would exhaust the resources found per patch during the search process. Varying magnitudes of patch types can include variability of reward gains per resource found and size of patch (Charnov, 1976). One can think of a classic example such as foraging for berries and the berry bushes being the resource patch. The varying sizes of bushes become indicators of how many berries a forager can extract along with varying quality of berries in that bush independent of bush size (Krebs, Kacelnik, Taylor, 1978; Pyke, 1984). For example, a forager can encounter a large bush with many berries and with berries that are at the peak ripeness to replenish maximum nutritional value per berry consumed. On the other hand, a forager can also encounter a small bush with berries that are of poor quality, meaning that the amount of nutritional value or quality of energy gained per consumption of berry is low. In an environment of depleting resources and limited energy, the problem of foraging is centered around when the best time to move onto a new resource patch is, as to not have a detrimental rate of energy intake per foraging bout (Pyke, Pulliam, & Charnov, 1977).

For a predator to be optimal in its foraging it needs to balance its energy intake over the time spent in a patch. Depleting resources will make it so that the time between successive consumption increases over time. Going back to the berry bush example, one can imagine foraging in a large berry bush and being pleased at the number of berries easily available to you. At some point, finding the rest of the berries in the bush will become more labor intensive and you will be putting too much energy into trying to fully deplete that resource patch. If a forager stays too long in a patch, then the energy intake rate will be too low for the amount of energy spent. The optimization of this behavior is therefore outlined by the Marginal Value Theorem, which states that energy intake ($g_j(T_j)$) over time spent in a patch (T_j) is equal to the average capture rate in the habitat ($En*$).

$$(1) \quad \frac{\partial g_j(T_j)}{\partial T_j} = En*$$

This is optimal under several more assumptions, which include a constant between patch travel cost and random encounter of patch types. This means that the length of time of traveling between patches should be independent of time spent in a patch, but the time spent in a patch will depend on the time traveling between patches. The original model holds that parameter constant to highlight time allocation in a patch without complicating the model to need to explain a rational choice analysis when traveling to different patch types (Pyke, 1984). This

simplification allows the theory to focus on what an optimal forager is and how their behaviors are demonstrated to be optimal on average. Natural selection is used to explain how optimal foragers are formed to follow optimal behavior as outlined by the MVT. The caveat to this is that foragers have access to the complete information of their habitats in order to make the model work (Kamil & Yoerg, 1982; Waddington, 1985). The limitations to this definition are discussed further in the paper but for now we review the original assumptions of the model and theory.

A further breakdown of the model is needed to fully understand the theoretical bounds and epistemic purpose of using the MVT to describe optimal foraging behavior. The first parameter in the model (P_i) takes into account the proportion of the visited patches that are of type i. As previously mentioned, types of patches in this context can vary the amount of resources contained along with the type of quality of resources the patch holds. The second parameter (E_T) is used to set the energy cost of traveling between patches per unit time. The theory holds this value constant. Third, the parameter (E_{si}) determines how much energy to charge in a patch of type i per unit time. Given these fundamental parameters, it is possible to define how the relationships in the model are built. The gained energy from foraging for any given number of time units (T) in a patch type of i minus all energy costs that do not include the cost of foraging in a patch is denoted as $h_i(T)$. When correcting for the energy cost of searching in a patch, the function is denoted as:

$$(2) \quad g_i(T) = h_i(T) - E_{si} \cdot T$$

Next, the relationship that describes the average time spent in a patch (T_u) includes the time it takes to travel between patches t plus the time spent in patches.

$$(3) \quad T_u = t + \sum P_i \cdot T_i$$

Again, the variability of values that can come from different patch types is denoted by i and similarly it can be shown how the average energy from a patch (E_e) is described.

$$(4) \quad E_e = \sum P_i \cdot g_i(T_i)$$

Given the description above, we can now present the relationship of the net energy intake rate (E_n).

$$(5) \quad En = \frac{E_e - t \cdot E_T}{T_u}$$

Equation (5) is fundamentally an energy balancing function and with the equations presented above it becomes clear how the model acquires the values now needed to make a case for optimal behavior.

The time spent traveling between patches t is argued to be independent of the time spent in any one patch T_i , but the time spent in any one patch cannot be said to be independent of the time spent traveling between patches. For example, a forager can spend a large amount of time traveling between berry bushes but once it arrives to the bush, factors such as patch size and resource quality will determine how long a forager stays (McNamara & Houston, 1985). On the other hand, if a forager stays either too long or short in one patch, the decision to travel more or

less between patches now becomes important for its survival. If an animal spends too much energy in one bush, then it will need to replenish its energy as soon as possible and a long between patch travel time would not be the most attractive choice. Given this theoretical grounding, the parameters for how to obtain the values of importance can now be compared to set what is optimal in the theory. Equation (5) is written in form (6) to equate with (7) which is now an equation that allows us to look at the process from a single patch (j) perspective.

$$(6) \quad En = \frac{\sum P_i \cdot g_i(T_i) - t \cdot E_T}{t + \sum P_i \cdot T_i}$$

$$(7) \quad En = \frac{P_j \cdot g_j(T_j) + A}{P_j \cdot T_j + B}$$

Values A & B are the result of equating 6 & 7 and are not functions of T_j in which the MVT states that a forager only has control over T_j when patch j is visited. Finally, for all $i = j$ given that all T_i are at their optimal values, En is written as $En *$. We then are able to describe it as equation (1), where the $En *$ average capture rate in the habitat is the optimal threshold for when to leave any patch j (Charnov, 1976).

The original model as described above faced some limitations when trying to account for the complete information an optimal forager may lack and variability of patch quality. High variability across patches makes it difficult for a forager to follow a generalized optimal policy if encounters are random and resources non-renewable. Furthermore, behavioral ecologists have not been able to mend the problem of incomplete information in food-finding behavior despite attempts to model patch sampling as a means for describing information gathering (Stephens & Krebs, 1986). Markov chain models have been used to describe patch sampling, in which an optimal forager who faces uncertainty surveys patch quality (Bobisud & Potraz, 1976; Stephens, Charnov, 1982). Modeling an optimal forager with incomplete information of their habitat has shown that it is optimal to stay longer in a patch past the optimal policy (Oaten, 1977; Green, 1980). Staying longer allows the forager to learn valuable information about their habitat such as patch type distributions and variability of reward gains (Pyke, 1984). The limitations of the original optimal forager with complete information is reviewed with the extension to cognitive processes in the next sections.

OFT in Cognition

Since the implementation of the MVT to concisely describe optimal foraging in food finding behavior, proponents of OFT have extended how the model can describe cognitive processes framed as foraging. This review highlights relevant historic publications and will not serve as an extensive list of all instances of MVT application in human cognition. One of the first examples was with the declaration of treating information spaces and their human-computer interactions as a foraging process (Pirolli & Card, 1995). Information Foraging (IF) then became the study of the trade offs that people consider in information seeking tasks, such as trying to balance the quality of information gain versus the cost of performing an activity in interaction with the design of the information space. Furthermore, the design of the information space was more commonly treated as an aid in not just the storage and handling of the information but also

with the optimization of its access. Examples of this in the IF theory tested spaces anywhere from rank order search results from search engines, citation links in papers, and electronic mail to more physically tangible spaces such as office workspaces. The theory also states that human computer interaction adds another dimension to information foraging. That is to say that people put in effort to enrich the foraging process because of the ability to design information spaces (Pirolli & Card, 1999).

A key distinction from the classic food finding forager in IF is the ability for a forager to design their information spaces which involves a focus on rational thought. For our purposes, rational thought is used in connection to the ability to design and enrich our information spaces for the purposes of optimizing the gains of foraging behavior. Something that an optimal forager from a behaviorist perspective can offload to natural selection and not to internal cognitive processes without having to worry about modeling decision making. IF theory puts its own model of higher-level cognition (ACT-IF), inspired by Anderson (1996) and the Adaptive Control of Thought – Rational (ACT-R) model for higher level cognition as a computational production system. Wayfinding and general knowledge in information spaces are described by scent detection processes, while the tradeoff dilemma of maximizing gains in current information patches is implemented by the MVT. On top of these processes, the relationship of an information forager who is capable of habitat design and purposeful foraging enrichment is explained by the ACT-IF model. A further break down of the ACT-IF model can be seen in Pirolli & Card (1999). In this section we focus on the extension of MVT implementation when framed for cognitive processes.

A more direct connection to behavioral ecology and foraging in humans came about with the testing of adaptability in foraging strategies. In reference to the classic OFT food finding behavior, an animal is adapted to their specific habitat and the knowledge needed to perform optimally over time is grounded in natural selection (Charnov, 1976; Pyke et al., 1977). The original theory did not present a description of foraging as a cognitive process but rather focused on the generalization of the behavior. Whether an animal had to make a rational choice or care about adapting its strategy to fit unknown habitats was not the original motivation for OFT. In the case of studying humans as an expert generalist, the question of adaptability in strategies became relevant in the connection to the behavioral ecology literature. Hutchinson, Wilke & Todd (2008) tackled this issue by placing participants in a fishing computer game. In the game, participants were given up to 45 mins to catch as many fish as they could and were monetarily incentivized to catch more fish. A resource patch in the game was implemented as a pond and the participant was allowed to switch to a new pond at any time without being able to run out of ponds to switch to. The conditions in the experiment varied fish across ponds from even, Poisson, and aggregated distributions. A second version of the aggregated condition made between patch travel time longer from 15 secs to 25 secs. The optimal policy for when to switch to a new patch across varying distributions is only dependent on the number of items caught in a patch N and the time spent in a patch T (Iwasa, Higashi, & Yamamura, 1981).

Theoretical predictions for optimal policies were tested through simulations to compare how participants differ from optimal. If a resource distribution is set to even, then the optimal

policy is a 'fixed-N-rule', in which participants should leave to a new patch after finding a certain number of items. A Poisson distribution across patches has a 'fixed-T-rule', describing that a forager needs switch at a fixed time. In an aggregated distribution the optimal policy should be defined by the 'incremental rule', in which a forager extends the time in a patch for every new item found. Conversely, if the distribution is sparse then the forager should follow a 'decremental rule' and decrease the amount of time in a patch for every item found. The human data collected showed that people did not follow the optimal policies across conditions. Furthermore, a performance comparison with a 'giving-up time' heuristic was done to compare human performance with a general heuristic. In this comparison, the performance was closer than the optimal policy but still statistically different. For all conditions, people demonstrated linear positive correlations for time spent per item found in a patch. A Cox Proportional Hazards Regression model was used to determine that time between item captures was the biggest predictor for switching to a new patch. It was concluded that equating sufficient knowledge in the environment of the game was not easy and that participants were only basing their decisions from the success of their last two captures. The MVT uses the optimal forager theory to explain the robust knowledge a forager has from an environment that they developed in, which participants did not have. Therefore, a possible explanation for the data is that participants were defaulting to a heuristic as a failsafe on their performance instead of risking the confirmation of an optimal policy.

Following the previous study's lead, a new study sought to measure internal word search with given sequences of letters as the resource patch. A group of German speakers were given predefined letter sequences and told to create as many words as they can with that sequence, much like a game of Scrabble. Participants were monetarily incentivized to create more words and could switch to a new 'patch' of letter sequences whenever they deemed necessary but the movement to the next patch had a time cost. The study had limitations for defining specific resource distributions across patches, so they simplified the method to only look at aggregated (high variance) and dispersed distributions (low variance). A letter's probability of use in the German language was used to fit them into higher or lower variance distributions in the sequences. Theoretical predictions posed that dispersed distributions would have the 'decremental rule' as the optimal policy while aggregated distributions should follow the 'incremental rule'. Results replicated the Hutchinson et al., (2008) study and showed that the interval time of the last two words found was the strongest predictor for when to switch to a new patch of letters instead of the total number of solutions found over the total time spent (Wilke, Hutchinson, Todd, & Czienskowski, 2009).

Again, there are some limitations to this kind of experimentation because of how difficult it is to put a person into the shoes of an optimal forager while also trying to avoid an a priori bias to make the participant produce ideal data. The participant was seen to perform a general heuristic based on short-term memory of performance. The authors raise some speculations for why participants stick to a short-term memory based decision making strategy. One of which explains that aggregated distributions are the most common in nature and that when faced with uncertainty, a participant purposely stays longer in a patch for confirmation of what the hidden distribution might be. Another speculation they give is that a general heuristic allows for

flexibility in learning of distribution types, while achieving personally satisfactory performance. In other words, a participant may be focused on trying to have a performance that gives them high reward. They are also new to the task and feel pressured to leave the experiment with some money instead of little to none. Since they are put on the spot to problem solve, they default to a strategy that will balance out monetary reward and information for the task. One can imagine that a participant who is determined to learn what hidden distributions exist, will either encounter it by chance or will have to learn it at a cost that they may not be prepared to risk. We argue that this level of uncertainty directly shapes the decisions of the participants. This argument will be further developed in the next sections, but we start off by introducing the extent to which these experimental methods capture optimal foraging behavior.

Internal search is further explored through semantic foraging. Hills, Jones, and Todd (2012) first explore this by analyzing how semantic memory retrieval of words during a fluency task reflect whether people follow optimal MVT guidelines. Patches in this internal semantic-memory landscape are the subcategories in which words tend to cluster. A prior study has demonstrated that fluency performance has two important components, which are clustering and switching (Troyer, Moscovitch, & Winocur, 1997). Switching from word clusters is framed as the giving-up event in patch foraging. Hills et al., (2012) study two possible ways in which patch boundaries may be defined. First, the static patch model defines a switch from a patch if the next word does not share a semantic subcategory from the word before the current word. Second, the fluid patch model defines a switch solely based on whether the previous word shares a subcategory with the current one. The static patch model assumes a scenario in which a person searches memory by first picking a subcategory and then exploiting it until they switch into another subcategory to exploit. For example, a participant would start by picking a subcategory such as pets and continue to exhaust their memory of pets until they decide to switch into another subcategory. The fluid patch model assumes that patch boundaries are more related to similarity of the previously retrieved word. An example of this would be retrieving pet category words such as 'dog', 'cat' and making the next word 'lion'. The words 'cat' and 'lion' would be part of the fluid patch and have a boundary between 'dog' and 'cat', while maintaining the pets as a patch. Results from this study favored the fluid patch model when comparing it to randomized data in which participants produced less fluid patch switches than the static patch model predicted. Some switches also fell into the class of being static and fluid, but those switches which were exclusively static did not take any longer to produce than non-switches. Between patch travel is predicted to have significantly longer time than within patch travel. Inter-item retrieval time demonstrated that people stay under their global averages until they move between patches, which follows optimal MVT strategies. Cosine similarity also demonstrated that words were less similar across fluid patches than within. A word's frequency was also shown to be the highest at the beginning of a patch and decreases as exploitation of the patch reaches its end. The word frequency feature has subset involvement of other lexical properties that form a spatial semantic space used to define context use of words involved (Adelman, Brown, & Quesada, 2006; Griffiths, Steyvers, & Tenenbaum, 2007). Despite word frequency being a coarse measurement, this adds evidence that word patches are not based on random attributes for retrieval.

The question of what exactly defines a patch in a person's memory during an internal semantic foraging task is further studied through the comparison of possible patch models. The favorability of the Troyer et al., (1997) based fluid-patch model was able to demonstrate the non-randomness of patch definition when tested to a static-patch model based on semantic subcategorization (Hills et al., 2012). This finding helped pose the question of whether a patch during the internal semantic foraging process is favored by associative or categorical attributes. A combination of an augmented Troyer et al., (1997) semantic scheme is used alongside with the BEAGLE model of semantic representations (Jones, Kintsch, & Mewhort, 2006; Jones & Mewhort, 2007) that was used previously (Hills et al., 2012) to reference what can constitute a patch whether it be associative or categorical. The associative model predicts that patch boundaries occur based on a continuous chain of word associations as opposed to a predefined category which is searched until exhausted. An example of a categorical based search process can be the finding of a category such as 'pets' and then consequential production of words that are directly under that category until a new category such as 'zoo animals' is accessed, determining a new patch switch. On the other hand, an associative based search can start with 'cat' and then proceed to associate with 'dog' while maintaining a chained relationship if the next word is 'wolf', in which both the first and last word of that sequence are joined by their connection to 'dog'. This allows for the 'pets' category and 'canine' categories to be part of one patch. Results point to the associative patch model to be the best representation in the semantic foraging data. The categorical patch model was seen to fit human data at times, but the associative model fit best for within patch and between patch data. One key finding relevant to the argument at hand was that there were no significant differences in performance between the two patch models (Hills, Todd, Lazer, Redish, & Couzin, 2015). Meaning that participants had another reason for favoring the associative model when performance, high or low, could be achieved with either search strategy.

Discussion from the results presented in (Hills et al., 2015) characterize the search process as semantic network foraging. Interestingly enough, random walks on the (Hills et al., 2012) data have shown to produce behavior akin to the optimal policies set by the MVT (Abbott, Austerweil, & Griffiths, 2015). In reference to previous discussion, we can now see that semantic relationships as stored in our memory are not random. They are maintained by rich contextual features of semantic relationships, but the behavior of accessing that information in our memory can be easily described by a random walk. Moreover, authors at the center of expanding the MVT guidelines to cognitive processes also point to related work such as in vision research and social systems to be framed under the explore & exploit dilemma shared by these processes (Hills, Todd, & Jones, 2015). Across these foraging modalities, there are implementations of probabilistic models contrasting the original deterministic MVT. Similarly, foraging work outside of the patch foraging framing in OFT have modeled foraging trajectories through probabilistic models such as Lévy Walks (Viswanathan et al., 1996, 1999). Research in area-restricted search has been able to mend points from patch foraging and Lévy Walk models to demonstrate both the randomness of the foraging behavior and the connection to the patchy distribution of resources that exists in the natural world (Kerster, Rhodes, & Kello, 2016). It may seem counterintuitive that a process that navigates complexities of the world at multiple levels of space and time can be best characterized as a low memory and random behavior. Literature in

philosophy and psychology is discussed in the next section to tackle this inquiry formally and a heuristic perspective on the Optimal Forager is defined afterwards to formalize the work reviewed.

Bounded Rationality

Probabilistic models have been used to handle some uncertainty in the space of an agent. As we have briefly stated before, many foraging domains have implemented a variation of a probabilistic model for the same reason. Such an implementation contrasts the original theoretical assumptions and epistemic orientation of the MVT in OFT. The primary view presented here will be Bounded Rationality as originally defined by Herbert Simon (1972; 1990; 1997) and extend by Gigerenzer & Selten (2002). Bounded Rationality takes a perspective on how animals make decisions and includes domain specific motivations along with environmental pressures put onto the animal. Gigerenzer et al., (1999) describes three main features behind the importance of Bounded Rationality. The first is psychological plausibility, which explains the importance of modeling rationality based on the emotional, social, behavioral, and cognitive features of a species. Next, domain specificity describes context sensitive heuristics an animal performs as opposed to having a strategy that is domain general across contexts. Lastly, ecological rationality pairs up the functional relationship of an animal's environment to their heuristics. Proponents of Bounded Rationality step away from looking at rationality as omniscient, optimal, and consistent. In turn, this perspective embraces a probabilistic nature to decision making but with a focus on a heuristic based approach.

Environmental pressures bound your decision making, and in many foraging cases time or energy become the driving factor for when to leave or further exploit a patch. An animal foraging about trying to survive will risk too much trying to find the optimal path, especially in highly connected spaces. Authors we have reviewed often make the connection to the dark forest metaphor as a thought experiment for foraging and uncertainties of the space. Equipped with a handy flashlight as the connection to our bounded perception, a person's primary goal when stuck in a dark forest is to make it out alive. The case of optimality goes out the window unless the forager was ready to face dire consequences of time and energy constraints in the dark forest. Whether you are looking for food, writing a manuscript, or looking for a partner you will be ultimately bounded by time restrictions and the energy available to you. This is why proponents of Bounded Rationality describe the decision making of animals as being fast & frugal. As previously discussed, we know that food finding behavior is fast adapting because it sits at the base of survival and reproduction. Another way to put it is that animals had to be fast and frugal about their food choices and stick to heuristics that gave them good results, but not worry about choices that would give them the absolute best results. This is done best by maximizing environmental affordances, which can be best accessed through the specialization of a species. Also, general heuristics are uniquely tuned to the individual level based on the satisficing principle (Ward, 1992). Going back to the first feature of bounded rationality, the unique parameters of an individual will guide the expression of a general heuristic. Although some species may be both considered apex predators, they will also have unique considerations for

reaching that goal (Simon 1955, 1956, 1957). Economic models have also shown the implementation for how satisficing can be used to explain market performance (Radner, 1975).

The Bounded Rationality approach not only differentiated itself from optimal-deterministic implementations but also from Bayesian models that required 'demonic' strength. This term was expressed to contrast the aspiration of having a unified formula of rationality, in which the philosopher Leibniz reportedly received from God in a written book of nature. The book itself was written in the language of Universal Characteristic in which all reasoning would be replaced with one calculus (Leibniz, 1951). The criticism such a proposal faced is in reference to Laplace's demon, in which the mathematician Pierre-Simon Laplace proposed the possibility of writing an equation that describes all past states and predicts all future states of the universe, with the caveat of needing to know all information about the universe at the given moment. The possibility for that universal equation is characterized as requiring 'demonic' strength. Therefore, Bayesian models fall under the unbounded rationality category and the MVT is categorized as optimization under constraint. Both of which are models aided by 'demons' and not faithful in their normative power. A study demonstrated that a 'Take the Best' heuristic outperformed the demonic multiple regression models in predicting cases dealing with homelessness in the U.S, inner city high-school drop out rates, professor salaries, and house prices among other examples (Czerlinski, Gigerenzer, & Goldstein, 1999). Furthermore, Gigerenzer & Selten (2002) point out that good old fashion artificial intelligence also demonstrates the lack of environmental rationality with the creation of robots that move two feet and wait forever trying to calculate the optimal direction to move next (Russell, 1997).

Gigerenzer (2019) further formally addressed the notion of rationality being bounded and the role of heuristics to extend axiomatic rationality. The main argument made is that axiomatic rationality works for small worlds. This terminology is in reference to Savage (1951) where an exhaustive set of future states are mutually exclusive along with a set of consequences that are exhaustive and mutually exclusive of one's actions per every single state; defining a small world. Conversely, Savage presents two examples where the maximization of expected utility in rationality is encountered such as in when playing chess or planning a picnic. When playing chess, there is a case of computational intractability because no human or machine (to our knowledge) can determine the exhaustive set of all possible states. For scale, the number of possible unique sequences (10^{120}) in the game are greater than the predicted number of atoms in the universe. The class of computationally intractable problems fall under NP-Hard, NP-Complete or somewhere in between. In Computer Science, these problems are generally defined as problems in which there are no algorithms that can solve and verify the solution efficiently in polynomial time. Other modern examples of computationally intractable problems are found in Nintendo games such as Super Mario (Aloupis, Demaine, Guo, & Viglietta, 2015). On the other hand, planning a picnic is defined as uncertain because all the states of consequences are not guaranteed to be mutually exclusive or exhaustive. In the words of the popular music group OUTKAST 'You can plan a pretty picnic, but you can't predict weather…'. When decision making is set in an intractable or uncertain space, bounded rationality along with a fast and frugal approach to problem solving is seen. Therefore, leaving the normative power of axiomatic

rationality limited to small world problems, where it is possible to know all states and properly define the probability of consequences.

This position purposely describes rationality as being bounded instead of limited because of the importance put on environmental pressures shaping decision making. For example, a person is capable of trying to find the optimal solutions in a task. In general, people feel special because of the ability and affordance we have to pursue such an endeavor of deep thinking, which is often used to differentiate us from the rest of the animal kingdom. On the other hand, Bounded Rationality points out that we more often rely and perform domain specific heuristics that put us on cruise control as we go about our day due to our expertise of our habitat (Czerlinski et al., 1999; Gigerenzer & Goldstein, 1999; Goldstein & Gigerenzer, 2002, 2009).

Heuristic Forager

So far, we have reviewed the original MVT model and the extent to which the classic OFT has been expanded into describing cognitive processes. The original notion of an Optimal Forager is defined by three key features. First, the development of optimal foraging behaviors is backed up through natural selection and specialization of species. Second, access to complete information, as necessitated by the original MVT model, is assumed to either come from the specialization of species adaptation or to be irrelevant because it is washed out on average when the habitat variability is low. Lastly, an optimal forager's habitat is assumed to have resources that are found in patches, non-renewable, have random encounters, and implicitly that variability of patch type distributions is not problematic. These assumptions work well to define a generalizable foraging behavior but nonetheless have limitations in their assumptions. What has been historically termed as the 'failure' of the Marginal Value Theory is related to problems of variability of patch types and incomplete information of a forager (Stephens & Krebs, 1986). Original motivations of the OFT use of MVT came from a behaviorist perspective. This means that the purpose of this theory was to unify a commonly observed animal behavior. To do that, researchers formed an epistemic boundary on internal processes. Like all models, the MVT had limitations in order to focus on the epistemic needs of traditional OFT. In fact, it was the success of the model that allowed researchers to ask more questions that touched upon cognitive processes. Research in patch sampling and attempts to mend the incomplete information issue were steered into dealing with what was happening inside the mind of a forager. When research had trouble fitting the data exactly to the model, proponents of the MVT pointed to incomplete information (Pyke et al., 1977; Werner, Mittelbach, & Hall, 1981). This started framing a scenario in which foragers had to behave optimally under several uncertainties. In order to navigate these low information scenarios, researchers now had to think of how to model rational choice and probabilities of expectation in foragers (Green, 1987; McNamara & Houston, 1985). Therefore, a cognitive perspective on foraging agents needs to rethink how an optimal forager is defined differently than with the original assumptions.

In the cases where participants where challenged with varying patch type distributions, the optimal policy as set by the MVT were not adhered to (Hills, Todd, & Jones, 2015; Hutchinson, Wilke, & Todd, 2008; Wilke et al., 2009). Furthermore, patches as relating to information spaces, memory, and semantics did not have simple patch boundaries as in the

behavioral ecology literature (Hills, Todd, & Jones, 2015; P. Pirolli, 2005; P. Pirolli & Card, 1999). In sum, a cognitive perspective has shown that the habitats we forage in are not simple and that foragers constantly account for uncertainty in their habitats. IF theory relates the capability to enrich our foraging processes by optimizing the design in our information spaces (Pirolli, 2007). Semantic foraging demonstrated that uncertainty of patch boundaries and of patch distributions for a forager come from the multilevel network connections that contextual semantic features have in language (Hills et al., 2012; Hills, Todd, & Jones, 2015). Foraging has also been seen to happen at different time-scales and across cognitive modalities (Hills, Todd, Lazer, et al., 2015). With this we propose the 'Heuristic Forager' to describe the features that a forager needs to have from a cognitive perspective.

The 'Heuristic Forager' (HF) will be defined by three key features that touch upon adaptation, information inference, and a functional relationship to its habitat. The original MVT referenced work stating that foraging behavior was fast adapting (Pyke et al., 1977). Behaviors directly related to eating were ones that animals had to prioritize in order to stay alive. Therefore, animals had to have a fast way of adapting to variability of the land in order to find food. Furthermore, the specialization of animals in unique habitats is referenced to point out that animals are not simply surviving but they are experts of their food consumption (Charnov & Schaffer, 1973; MacArthur, Diamond, & Karr, 1972; Pulliam, 1974; Schoener, 1969, 1971). With this, natural selection is used to defend the simplicity of the MVT and assume that animals have sufficient information to carry out optimal behavior. The HF definition is used to mend the problems the original assumptions had with claims of sufficiency. Both the fast adapting foraging behavior and specialization can be sufficient if we also turn our attention to the mind of a forager. The claim here is that a forager has to constantly make inferences about their environment, giving us room to expand the qualities of a foraging habitat which were simplified originally. A HF is therefore always assumed to not have complete information and to maximize performance under limited knowledge. Natural selection is now used to explain the development of domain specific heuristics instead of optimal.

In the same way that the MVT related the habitat's structure as shaping the functionality of optimal foraging policies, we also present the case that a complex and highly connected environment shapes the functionality for a HF. For a HF we define the habitat as being different than the simple patch space, and more akin to a network space. IF theory and semantic foraging have shown examples of cognitive processes navigating an explore & exploit dilemma on network spaces. Although the specifics to a network structure have not been tested for its implementation with the MVT, research has used network spaces to portray meaningfully connected complex spaces. A random walk on a network space has produced behavior that is in line with MVT (Abbott et al., 2015) but researchers have not implemented MVT guidelines to a forager over network spaces to measure its success. In IF theory, networks portray information spaces that would otherwise be difficult to access without enrichment to the process such as in the use of a hyperbolic connectivity tree. A study was conducted to compare performance on information access comparing the hyperbolic tree to a standard Windows file browser. The hyperbolic tree demonstrated better performance in tasks that had stronger cues but underperformed in comparison in tasks with less cues (Pirolli, Card, & Wege, 2003). A semantic

network is also used to simplify lists of features that are unknown and can simply be organized through association of use (Steyvers & Tenenbaum, 2005). For example, in semantic foraging known features are characterized by a small-world network connectivity. For every small-world network foraged in memory, there is also an edge of connectivity to another small-world network describing a not immediately visible feature (Abbott et al., 2015; Hills et al., 2012; Rhodes & Turvey, 2007). If our current semantic patch is being exploited by an association of farm animals, we might be able to stumble on a word such as 'chicken' and then start thinking of animals that we normally eat which would mean the jump to a new small-world network dictated by that current association. Therefore, the specialization of animal much in the way of Darwinian finches, is used to explain the access of information instead of implementing the need for complete information in a model.

Foraging in network spaces is used as a proxy in research and in our argument for the complex connectivity of associations in either semantic or informational space that may not be perceptually attainable to the forager. The dark forest metaphor is more applicable in connection to networks to explain how a searcher with a flashlight comes across new information. For every time the forager flashes their light in one direction, new paths in the forest become clear. This type of unknown complexity does not need to be modeled through a network space, but we stick to how previous research has portrayed it (Abbott et al., 2015; Hills et al., 2012; Hills, Todd, & Jones, 2015). Findings in both OFT and Lévy Walk literature have found their own way to describe foraging as a random process, because they are useful in handling uncertainties about space. This is true whether the connectivity of the space gives us a local vs global vision for search or if the distribution of the resources is unknown. (Frederic Bartumeus, M. G. E. Da Luz, 2005; Plank & James, 2008; Rhodes & Turvey, 2007; Robertson, Guckenheimer, Masnick, & Bacher, 2004; Stephens & Charnov, 1982; Viswanathan et al., 1996, 1999). In connection to Bounded Rationality theory, we claim that these probabilistic processes are signs of foragers performing heuristics to maximize their functional relationship with their environments. Therefore, making the features of the defined HF more relevant to the models used to implement foraging behavior than the original Optimal Forager definition.

Discussion

A HF with Bounded Rationality presents a case for accepting the domain specific features an animal has in relation to their own environment. This changes the script on adaptability and observed behaviors. We presented the limitations from the original perspective implemented through the MVT, which did not allow for looking through the perspective of the forager. When we start seeing through what a forager is thinking, we see that the forager has many things to keep track of. So many, that through its adaptation it has stuck to domain specific heuristics to guide it forward. On the other hand, a domain general heuristic detaches itself from adaptability although still not relying on 'demonic' strength for model implementations. A study reviewing scheduling problems such as the Traveling Salesman Problem found that ~84% of problems were intractable (Lawler, Lenstra, Rinnooy Kan, & Shmoys, 1993). Moreover, there is general argument in computer science that computationally intractable problems are the focus of interest be it via machine learning or neural networks (Tsotsos, 1991). Foraging as originally

described by OFT fit the problem as an Optimal Control Problem, but if a problem is computationally intractable, then such a classification no longer makes sense. Therefore, it is evident that cases of rationality relevant to us do not deal in a small world scenario. Meaning that most animals have had to learn to stick to simple, high performing heuristics such as observed through probabilistic models in foraging and not through the route that Optimal Control Problems take to solve tasks.

The low memory probabilistic models implementing random walks have been seen to fit data in various foraging scenarios. It can be easy to make the assumption that their success is evidence for a domain general approach. In fact, Hutchinson et al., (2008) compares performance of human and heuristic to support the robustness a domain general heuristic can have. This is in comparison to optimal predictions which require the full knowledge of resource distributions. Therefore, showing that a domain general heuristic can have satisfactory performance despite its detachment of domain specific factors. In a way, these domain general heuristics assume a high level of uncertainty and instead stick to a consistent and simple behavior. In other words, a blind forager has no need for a flashlight in a dark forest. Our argument presents these domain general heuristics as evidence for rethinking the modeling of searching behavior. If a blind behavior can show relevant performance in comparison to demonic models, then it is worth rethinking the normative strength of these models. We propose that a domain general heuristic is a starting point for crafting a domain specific heuristic. The idiomatic comparison thus characterizing demonic models as too hot and domain general heuristics as too cold. Putting forth a testable hypothesis that a properly captured domain specific behavior will have performance in between computational optimal and domain general. Our purpose here is to make an argument for modeling these behaviors with domain specific features that do not require demonic strength. An argument towards domain general over domain specific heuristics can be made, but for our purposes we suffice with demonstrating our inclination towards the comparison of demonic models and leave that debate to be flushed out in future discussions.

In sum, we have made a case to update the notion of an Optimal Forager to a Heuristic Forager in order to strengthen the normative power of foraging models when considering cognitive processes. At the core, our argument is not meant to be adversarial but rather encompassing the success of the MVT in OFT. With this, we would also update the notion of optimality instead of completely putting it aside. From an individual forager's perspective, we know that the foraging process is a means to an end. Findings in OFT work have even pointed out that additional increases in energy intake do not increase fitness (Krebs & McCleery, 1984). This combined with principles of satisficing and being fast & frugal in the Bounded Rationality perspective, shows us that it is often not the case for trying to optimally solve problems but rather to perform well and be adaptive. This does not mean that we should abstain from involving optimality but rather we can have a new perspective on where the optimality is. We conclude with framing foraging as an Optimal Control Problem at a different scale. This is in reference to the enrichment that happens in Information Foraging. For example, if it takes me 10 mins to find a folder in my informational space, then through enrichment I will make it so that the next time it should not take me as long (Pirolli & Card, 1999; Pirolli, 2007). Another more extreme example highlights the convenience of completely getting rid of the foraging process

when possible. New programs and services such as recommender systems have made it a business to cater our entertainment instead of allowing us to forage around flipping through channels (Resnick, Varian, 1997). Similarly, dogs have traded the need to hunt for food and instead remain loyal to their human owners who feed them in return (Frank & Frank, 1982). In both cases, a change in the relationship with the environment seeks to relieve one from spending energy in foraging and reaching a desired goal. Foraging then being the transient state of escaping the purgatory-esque chains of our basic needs. In order to transcend into affordance that shines a brighter light during our visits to the dark forest.

References

Abbott, J. T., Austerweil, J. L., & Griffiths, T. L. (2015). Optimal Foraging Random Walks on Semantic Networks Can Resemble Optimal Foraging. *Psychological Review*, *122*, 558–559. https://doi.org/10.1037/a0038693

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual Diversity, Not Word Frequency, Determines Word-Naming and Lexical Decision Times. *Psychological Science*, *17*(9), 814–823. https://doi.org/10.1111/j.1467-9280.2006.01787.x

Aloupis, G., Demaine, E. D., Guo, A., & Viglietta, G. (2015). Classic Nintendo games are (computationally) hard. *Theoretical Computer Science*, *586*, 135–160. https://doi.org/https://doi.org/10.1016/j.tcs.2015.02.037

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American psychologist*, *51*(4), 355.

Bobisud, L. E., & Potratz, C. J. (1976). One-trial versus multi-trial learning for a predator encountering a model-mimic system. *The American Naturalist*, *110*(971), 121-128.

Charnov, E.L. (1976). Optimal foraging theory: the marginal value theorem. *Theoretical Population Biology*, *9*, 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Charnov, Eric L. (1976). Optimal Foraging: Attack Strategy of a Mantid. *The American Naturalist*, *110*(971), 141–151. https://doi.org/10.1086/283054

Charnov, Eric L, & Schaffer, W. M. (1973). Life-History Consequences of Natural Selection: Cole's Result Revisited. *The American Naturalist*, *107*(958), 791–793. https://doi.org/10.1086/282877

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In *Simple heuristics that make us smart.* (pp. 97–118). New York, NY, US: Oxford University Press.

Eisenbach, M., & Lengeler, J. W. Chemotaxis. 2004. *London: Imperial College Press*, *16*(499), 1326.

Frank, H., & Frank, M. G. (1982). On the effects of domestication on canine social development and behavior. *Applied Animal Ethology*, *8*(6), 507–525. https://doi.org/https://doi.org/10.1016/0304-3762(82)90215-2

Frederic Bartumeus, M. G. E. Da Luz, G. M. V. and J. C. (2005). Animal Search Strategies : A Quantitative Random-Walk, *86*(11), 3078–3087.

Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart* (pp. 3-34). Oxford University Press.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The take the best heuristic. In *Simple heuristics that make us smart.* (pp. 75–95). New York, NY, US: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition

heuristic. *Psychological Review*. Goldstein, Daniel G.: Max Planck Inst for Human Development, Ctr for Adaptive Behavior & Cognition, Lentzeallee 94, Berlin, Germany, 14195, goldstein@mpib-berlin.mpg.de: American Psychological Association. https://doi.org/10.1037/0033-295X.109.1.75

Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making*, *62*, 88.

Gigerenzer, G. (2019). Axiomatic rationality and ecological rationality. *Synthese*, 1-18.

Goldstein, D. G., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting*, *25*(4), 760–772. https://doi.org/https://doi.org/10.1016/j.ijforecast.2009.05.010

Green, R. F. (1980). Bayesian birds: a simple example of Oaten's stochastic model of optimal foraging. *Theoretical Population Biology*, *18*(2), 244-256.

Green, R. F. (1987). Stochastic Models of Optimal Foraging BT - Foraging Behavior. In A. C. Kamil, J. R. Krebs, & H. R. Pulliam (Eds.) (pp. 273–302). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4613-1839-2_8

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211—244. https://doi.org/10.1037/0033-295x.114.2.211

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440. https://doi.org/10.1037/a0027373

Hills, T. T., & Dukas, R. (2012). The evolution of cognitive search. *Cognitive search: Evolution, algorithms, and the brain*, 11-24.

Hills, T. T., Todd, P. M., & Jones, M. N. (2015). Foraging in Semantic Fields: How We Search Through Memory. *Topics in Cognitive Science*, *7*(3), 513–534. https://doi.org/10.1111/tops.12151

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., & Couzin, I. D. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*(1), 46–54. https://doi.org/https://doi.org/10.1016/j.tics.2014.10.004

Hutchinson, J. M. C., Wilke, A., & Todd, P. M. (2008). Patch leaving in humans: can a generalist adapt its rules to dispersal of items across patches? *Animal Behaviour*, *75*(4), 1331–1349. https://doi.org/10.1016/j.anbehav.2007.09.006

Iwasa, Y., Higashi, M., & Yamamura, N. (1981). Prey distribution as a factor determining the choice of optimal foraging strategy. *The American Naturalist*, *117*(5), 710-723.

Jones, M. N., Kintsch, W., & Mewhort, D. J. (2006). High-dimensional semantic space accounts of priming. *Journal of memory and language*, *55*(4), 534-552.

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological review*, *114*(1), 1.

Kamil, A. C., & Yoerg, S. I. (1982). Learning and Foraging Behavior BT - Ontogeny. In P. P. G. Bateson & P. H. Klopfer (Eds.) (pp. 325–364). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-7578-8_7

Kerster, B. E., Rhodes, T., & Kello, C. T. (2016). Spatial memory in foraging games. *Cognition*, *148*, 85–96. https://doi.org/10.1016/j.cognition.2015.12.015

Kirk, D. E. (2012). *Optimal control theory: an introduction*. Courier Corporation.

Koshland, D. (1980). *Bacterial chemotaxis as a model behavioral system* (Vol. 2). Raven Pr.

Krebs, J. R., & McCleery, R. H. (1984). Optimization in behavioural ecology. Behavioural Ecology: An Evolutionary Approach.

Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., & Shmoys, D. B. B. T.-H. in O. R. and M. S. (1993). Chapter 9 Sequencing and scheduling: Algorithms and complexity. In *Logistics of Production and Inventory* (Vol. 4, pp. 445–522). Elsevier. https://doi.org/https://doi.org/10.1016/S0927-0507(05)80189-6

Leibniz, G. W. (1951). Selections.

MacArthur, R. H., Diamond, J. M., & Karr, J. R. (1972). Density Compensation in Island Faunas. *Ecology*, *53*(2), 330–342. https://doi.org/10.2307/1934090

McNamara, J. M., & Houston, A. I. (1985). Optimal foraging and learning. *Journal of Theoretical Biology*, *117*(2), 231–249. https://doi.org/https://doi.org/10.1016/S0022-5193(85)80219-8

Oaten, A. (1977). Optimal foraging in patches: a case for stochasticity. *Theoretical population biology*, *12*(3), 263-285.

Pirolli, P., & Card, S. (1995). Information foraging in information access environments. In *Chi* (Vol. 95, pp. 51-58).

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2003). The effects of information scent on visual search in the hyperbolic tree browser. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *10*(1), 20-53.

Pirolli, P. (2005). Rational Analyses of Information Foraging on the Web. *Cognitive Science*, *29*(3), 343–373. https://doi.org/10.1207/s15516709cog0000_20

Pirolli, P., & Card, S. (1999). Information Foraging. *Psychlogical Review*, (January), 643–675. https://doi.org/10.1037/0033-295x.106.4.643

Pirolli, P. L. T. (2007). *Information Foraging Theory: Adaptive Interaction with Information. Human Technology Interaction Series*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195173321.001.0001

Plank, M. J., & James, A. (2008). Optimal foraging: Lévy pattern or process? *Journal of the Royal Society Interface*, *5*(26), 1077–1086. https://doi.org/10.1098/rsif.2008.0006

Prokopy, R. J., & Roitberg, B. D. (1984). Foraging Behavior of True Fruit Flies: Concepts of

foraging can be used to determine how tephritids search for food, mates, and egg-laying sites and to help control these pests. *American Scientist*, *72*(1), 41–49. Retrieved from http://www.jstor.org/stable/27852437

Pulliam, H. R. (1974). On the Theory of Optimal Diets. *The American Naturalist*, *108*(959), 59–74. https://doi.org/10.1086/282885

Pyke, G H, Pulliam, H. R., & Charnov, E. L. (1977). Optimal Foraging: A Selective Review of Theory and Tests. *The Quarterly Review of Biology*, *52*(2), 137–154. https://doi.org/10.1086/409852

Pyke, Graham H. (1984). Optimal Foraging Theory: A Critical Review. *Annual Review of Ecology and Systematics*, *15*, 523–575. Retrieved from http://www.jstor.org/stable/2096959

Radner, R. (1975). Satisficing BT - Optimization Techniques IFIP Technical Conference: Novosibirsk, July 1–7, 1974. In G. I. Marchuk (Ed.) (pp. 252–263). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-38527-2_34

Raichlen, D. A., Wood, B. M., Gordon, A. D., Mabulla, A. Z. P., Marlowe, F. W., & Pontzer, H. (2014). Evidence of Lévy walk foraging patterns in human hunter–gatherers. *Proceedings of the National Academy of Sciences*, *111*(2), 728 LP – 733. https://doi.org/10.1073/pnas.1318616111

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56-59.

Resnikoff, H. L. (2012). *The illusion of reality*. Springer Science & Business Media.

Russell, S. J. (1997). Rationality and intelligence. *Artificial intelligence*, *94*(1-2), 57-77.

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A: Statistical Mechanics and Its Applications*, *385*(1), 255–260. https://doi.org/10.1016/j.physa.2007.07.001

Robertson, S. S., Guckenheimer, J., Masnick, A. M., & Bacher, L. F. (2004). The dynamics of infant visual foraging. *Developmental Science*, *7*(2), 194–200. https://doi.org/10.1111/j.1467-7687.2004.00338.x

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical association*, *46*(253), 55-67.

Schoener, T. W. (1969). Models of Optimal Size for Solitary Predators. *The American Naturalist*, *103*(931), 277–313. https://doi.org/10.1086/282602

Schoener, T. W. (1971). Theory of Feeding Strategies. *Annual Review of Ecology and Systematics*, *2*, 369–404. Retrieved from http://www.jstor.org/stable/2096934

Simon, Herbert A. (1955) "A behavioral model of rational choice", *The Quarterly Journal of Economics*, vol. 69, n. 1, February: 99-118, compiled in, and quoted from, Simon (1957: 241-260).

Simon, Herbert A. (1956) "Rational choice and the structure of the environment", *Psychological Review*, vol. 63, March, compiled in, and quoted from, de, Simon (1957: 261-273).

Simon, Herbert A. (1957) *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*, New York: John Wiley and Sons.

Simon, H. A. (1972). Theories of bounded rationality. *Decision and organization*, *1*(1), 161-176.

Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15-18). Palgrave Macmillan, London.

Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3). MIT press.

Stephens, D. W., & Charnov, E. L. (1982). Optimal foraging:some simple stochastic models. *Behavioral Ecology and Sociobiology*, *10*, 251–263. https://doi.org/10.1007/BF00302814

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton University Press.

Steyvers, M., & Tenenbaum, J. B. (2005). The Large-Scale Structure of Semantic Networks: Statistical Analyses and a Model of Semantic Growth. *Cognitive Science*, *29*(1), 41–78. https://doi.org/10.1207/s15516709cog2901_3

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why Is Infant Language Learning Facilitated by Parental Responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138–146.

Tsotsos, J. K. (1991). Computational resources do constrain behavior. *Behavioral and Brain Sciences*, *14*(3), 506–507. https://doi.org/DOI: 10.1017/S0140525X00071053

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., & Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature*, *381*(6581), 413–415. https://doi.org/10.1038/381413a0

Viswanathan, G. M., Buldyrev, S. V, Havlin, S., da Luz, M. G. E., Raposo, E. P., & Stanley, H. E. (1999). Optimizing the success of random searches. *Nature*, *401*, 911. Retrieved from http://dx.doi.org/10.1038/44831

Waddington, K. D. (1985). Cost-intake information used in foraging. *Journal of Insect Physiology*, *31*(11), 891–897. https://doi.org/https://doi.org/10.1016/0022-1910(85)90106-4

Ward, D. (1992). The Role of Satisficing in Foraging Theory. *Oikos*, *63*(2), 312–317. https://doi.org/10.2307/3545394

Weimerskirch, H., Salamolard, M., Sarrazin, F., & Jouventin, P. (1993). Foraging Strategy of Wandering Albatrosses Through The Breeding Season: A Study Using Satellite Telemetry. *The Auk*, *110*(2), 325–342. https://doi.org/10.1093/auk/110.2.325

Werner, E. E., Mittelbach, G. G., & Hall, D. J. (1981). The Role of Foraging Profitability and Experience in Habitat Use by the Bluegill Sunfish. *Ecology*, *62*(1), 116–125. https://doi.org/10.2307/1936675

Wilke, A., Hutchinson, J. M. C., Todd, P. M., & Czienskowski, U. (2009). Fishing for the right words: Decision rules for human foraging behavior in internal search tasks. *Cognitive Science*, *33*(3), 497–529. https://doi.org/10.1111/j.1551-6709.2009.01020.x