



Exploratory Search: Information Matters More than Primary Reward

Patrick Anselme

Department of Biopsychology, Ruhr-University Bochum

Email: Patrick.Anselme@rub.de

Citation – Anselme, P. (2023). Exploratory search: Information matters more than primary reward. *Animal Behavior and Cognition*, 10(4), 366-383. <https://doi.org/10.26451/abc.10.04.05.2023>

Abstract – In the study of animal foraging, resource exploitation (prey pursuit, handling, and consumption) has received much more attention than the search or exploratory process that leads predator to potential prey—whatever they are. Yet, in an unpredictable environment, exploration is crucial to optimize resource exploitation, or at least make this latter effective enough, and maintain organisms alive and capable of reproduction in the long term. I argue that environmental exploration requires psychological mechanisms that differ from those of resource exploitation. During exploration, organisms attempt to resolve the uncertainty about reward procurement rather than attempting to obtain reward. Behaviors that do not maximize reward procurement in some experimental designs are often described as suboptimal or even “irrational.” However, these designs might expose organisms to conditions that stimulate exploration more than exploitation. I suggest three general psychological principles assumed to govern environmental exploration, and report experimental evidence that justifies each of them. These principles may help account for behaviors difficult to explain by means of traditional theoretical frameworks.

Keywords – Cue, Reward, Effort, Uncertainty, Incentive, Information

Optimal foraging theory has been a revolution for the modern understanding of the rules that govern animal behavior, providing a powerful synthesis of findings from several research fields i.e., ecology, ethology, psychology, evolution, and microeconomy (e.g., Charnov, 1976; Pyke, 1984; Stephens & Krebs, 1986). The basic principle behind optimal foraging theory suggests that organisms always attempt to maximize reward rate per unit time—at least, in the long run. However, this theory describes idealized foragers deciding appropriate actions after computing the costs and benefits associated with prey-directed behaviors such as pursuit, handling, and consumption (Stephens & Krebs, 1986). Those idealized foragers possess complete knowledge about the availability of food items in a patch as well as about the location of and the distance from other patches. In other words, optimal foraging theory is focused on the question of resource exploitation and does not pay (much) attention to the question of environmental exploration—even though food search was early recognized as an essential part of foraging activity (MacArthur & Pianka, 1966). However, exploration is necessary to optimize resource exploitation as much as possible in a local environment, where interfering variables (predation, weather, heterogenous reward distribution, physical obstacles, etc.) limit reward procurement.

Exploratory search can be seen, in a large part, as an adaptive response to the interfering variables that prevent the optimal exploitation of a local environment. Although the payoffs of exploratory search are not guaranteed in advance and may even temporarily disadvantage immediate resource exploitation in a local environment, exploratory search is likely to maximize fitness otherwise. Let us illustrate this idea through an example. The individuals from species whose evolution produced “swarm” intelligence, like bees and ants, carry out highly specialized tasks that benefit collectivity. Among forager bees, some individuals display scouting behavior: they search for new nest sites and new food sites, even when highly

profitable patches have been found (Lindauer, 1952; Seeley, 1983). Food scouts represent 5-25% of a colony's foraging force and they do not collect food; they only explore the surroundings and communicate the new profitable locations to other forager bees—called recruits—through the waggle dance. Recruits are more numerous and follow the instructions provided by scouts to exploit a patch; they never produce scouting behavior (Liang et al., 2012). Such a dissociation between exploration and exploitation, magnified in Aculeate Hymenoptera (Jander, 1997), indicates that these two behavioral components correspond to different tasks, and hence depend on distinct biobehavioral mechanisms (Chatterjee et al., 2021; Cook et al., 2019; Huang et al., 2022; Liang et al., 2012). Both are necessary to ensure the effective functioning of a hive of dozens of thousands of individuals, but they are also necessary for a single individual working for its own survival. There is computational evidence that individual foragers can consume more food items and put on more fat reserves in an environment with food resources whose distribution is scarce rather than abundant, forcing them to explore a larger area (Anselme et al., 2017, 2018). The detection of food or food-predictive cues is not of primary importance in this process. A forager acting without landmarks (“myopic”) that eats only when encountering food while nutritionally depleted survives longer than a forager with perfect detection capability that always eats when food is encountered i.e., the wastage of food remains limited, and hence available longer, for the myopic forager (Rager et al., 2018). Thus, exploratory search is effective in improving fitness and should be dissociated from the optimal foraging theory's rule of thumb that reward rate per unit time must be maximized.

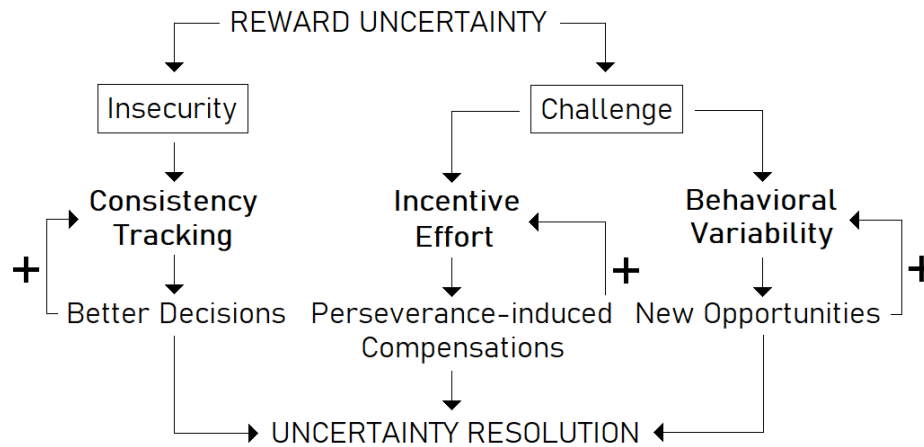
Exploration has traditionally been interpreted as a product of curiosity. For some theorists, curiosity is an aversive state aimed to reduce various forms of uncertainty—novelty, complexity, conflict—in the environment through information gathering (Berlyne, 1960; Loewenstein, 1994). For others, information is sought to downregulate arousal when too high (stress) but also to upregulate arousal when too low (boredom; Hebb, 1955; Sansone & Harackiewicz, 2000). In this paper, the role of curiosity in exploration is not denied and there are some connections between these views and the perspective developed here. However, curiosity is not viewed as a prerequisite for exploration, which has been identified in brainless organisms (e.g., McNickle et al., 2009; Yip et al., 2014). It is argued that unguaranteed reward (whatever the reason) leads to a perception of uncertainty as a challenge to be met. Although uncertainty is aversive, this perception is usually not a source of avoidance; organisms are rather willing to engage in the challenging situation. Organisms evolved to meet a number of challenges in their environment, and I show that this reaction to uncertainty is beneficial to health and hence to survival.

Resource exploitation can be seen as a local phenomenon that relates to reward consumption and the approach of reward-predictive cues that are detected or can be expected. By contrast, environmental exploration is more global and aims to determine where, when, and how to find those cues or rewards through information seeking. Exploration is an investment for the future, as it may have secondary benefits—such as optimizing safety in an open environment (Whishaw et al., 2006). However, if cues and rewards—or even hunger (e.g., Inglis et al., 1997)—do not motivate exploration, what other factors capable of optimizing fitness are behind it? These factors remain poorly understood (Bartumeus et al., 2016; Pisula, 2009).

In this paper, I define three principles assumed to underlie exploratory search. The first principle is called *consistency tracking* and means that organisms exposed to unguaranteed reward seek information about the consistency of cue-reward pairings more than the cues and the rewards per se. This process is a response to uncertainty induced reward insecurity, leading to better knowledge of the causal structure of the environment, and hence to better decisions. The second principle, *incentive effort*, posits that organisms exposed to unguaranteed reward strengthen and/or lengthen exploratory activity to compensate for their lack of cognitive control in the situation. Finally, the third principle, *behavioral variability*, suggests that organisms exposed to unguaranteed reward extend the spatial and/or temporal range of their responses because this strategy increases the opportunities for new successful encounters. Incentive effort and behavioral variability are assumed to result from the perception of uncertainty as a challenge to meet. These principles are not mutually exclusive and together act for uncertainty resolution (Figure 1). They may apply irrespective of the species and the neurobiological mechanisms involved. I provide compelling evidence that these three principles may provide a conceptual framework to better understand unresolved issues with respect to animal and human behaviors.

Figure 1

Principles Assumed to Promote Uncertainty Resolution Through Exploratory Search



Note. Principles assumed to promote uncertainty resolution through exploratory search, as well as their positive consequences (+) that make them both adaptive (at the species level) and reinforced (at the individual level). Reward uncertainty is a source of insecurity that favors consistency tracking but can also be perceived as a challenge, which leads to the expression of incentive effort and behavioral variability. All individuals are assumed to rank along the insecurity-challenge continuum. More details in the text.

Resource Exploitation and the Local Influence of Cues

The direct detection of an appetitive stimulus, or the expectation of one, leads to a non-random approach to the stimulus. In the words of Berridge (2007), incentive salience (or “wanting”) is the process that “transforms the brain’s representation from a mere perception or memory into a motivationally potent incentive” (p. 409). The attribution of incentive salience to a stimulus (perceived or recalled) therefore makes it appetizing and approached. This stimulus can be a reward or a cue predictive of reward delivery. This process is mainly dependent on the release of dopamine in the ventral and the dorsal striatum (e.g., DiFeliceantonio et al., 2012; Palmiter, 2008; Robinson & Berridge, 2013; Saunders & Robinson, 2012). In the Pavlovian context of repeated cue-food pairing, where the brief presentation of a cue is automatically followed by food delivery (autoshaping procedure), the cue is learned as a predictor of the upcoming food item. However, in some individuals, the incentive salience of the food is also gradually transferred—through dopaminergic signaling—to the cue and makes the cue attractive as well. These individuals are called sign-trackers, in opposition to goal-tracker individuals which do not develop any cue attraction—of note, a third category of individuals show intermediate performance. Although there are alternative interpretations of dopamine’s role (e.g., Beeler & Mourra, 2018; Salamone & Correa, 2002; Schultz, 1998), the body of evidence suggesting that dopamine is the major component of the motivational response to incentive stimuli is overwhelming.

Here is not the place to discuss incentive salience theory, but rather to emphasize that this theory is quite in agreement with the views that reward procurement is the survival-related factor to be optimized by organisms: a cue predictive of higher food quality/amount should be approached faster and preferred to a cue predictive of lower food quality/amount. Similarly, in an instrumental context, where the opportunity to obtain a reward is consequential to action, animals work more for a higher food quality/amount and the matching law predicts that organisms exposed to a choice will distribute their responses in proportion to the reinforcement rates available (Baum, 1974; Herrnstein, 1961). There is no doubt that cue-outcome and learned action-outcome associations can drive behavior. My point is that their influence is relatively limited, in the sense that cues and knowledge are mostly important for the exploitation of well-known locations. Cues and knowledge are less likely to explain behavior at a larger scale (Bartumeus et al., 2016; Humphries et al., 2012).

The classical distinction between focal and general search (Bond, 1983; Timberlake, 1994) goes to this direction because it implicitly suggests that two distinct processes are required for exploitation and exploration. However, such a distinction remains elusive regarding important points: What is exactly searched in one or the other case? If both search types are directed to cues and rewards, how is general search different from focal search? And how do organisms decide to switch from one to the other strategy? Incentive salience theory and the matching law are compelling explanations of focal search, where organisms try to reach a detected or expected stimulus. But general search seems different; it is not a strategy aimed to approach a stimulus. General search is rather an attempt to counter the potentially harmful effects of reward uncertainty or scarcity through exploration (Anselme, 2021; Anselme & Güntürkün, 2019). Understanding exploration requires a motivational theory that differs from incentive salience, that is, a theory in which search behavior is about information rather than about the immediate exploitation of cues and rewards.

In the next sections, I elaborate on the three principles presented in the introduction. It is shown how they contribute to the extraction of information by animals exposed to an environment with unpredictable food resources i.e., uncertainty-induced reduction in food access, in space and/or in time. Unpredictability in the environment can take many forms, but they all imply that organisms are exposed to significant events (or produce actions) with uncertain consequences (e.g., sometimes they are associated with food, sometimes they are not).

Consistency tracking

The principle of *consistency tracking* may sound counterintuitive: If the function of exploration is to find out reward sources, such as food, why not just say that organisms look for rewards and their associated cues rather than the consistency of their pairing. As mentioned earlier, brief occurrences of resource exploitation within an exploratory bout—a food item found accidentally may be consumed—should not lead to the conclusion that exploration is just exploitation, expressed at a larger scale. Here, exploration is presented as a search for information about conditioned and unconditioned stimuli—CSs and USs, respectively. Consistency relates to the concept of information in a restricted sense (see also the last section): under reward uncertainty, organisms primarily track the direct and indirect relations between stimuli and outcomes to be able to learn and use the structure of the environment they can then exploit. This process does not involve any form of consciousness. In other words, they attempt to determine *where, when, and how to find associations that predict as consistently as possible the presence of rewards*. In the absence of consistency tracking, survival would only be subject to the vagaries of positive and negative encounters. According to Keller et al. (2020), “[t]he information sought might involve specific environment landmarks (landmark information), how these landmarks relate to one another (configural information), whether the landmarks or structures together form an overall pattern (geometric information), or a combination of these information types” (p. 3). Of course, the information in question is rarely perfect in a natural setting because the consistency of cue-reward pairings is often partial and temporary. After following some signals, a squirrel may arrive under a walnut tree where nutshells stand on the ground, but this opportunity is transient and some of shells may be empty because the fruit decayed or was already eaten. Organisms can only attempt to use an information as accurate as possible about profitable places, times, and strategies. Consistency tracking is therefore reinforced (and adaptive), because learning the relation between events in an environment maximizes good decisions while minimizing bad decisions in this environment. Thus, consistency tracking participates to uncertainty resolution. This principle explains why exploratory search can occur in the absence of reward expectation in a novel environment (where possible cue-reward consistencies might be found) but is not shown in a familiar environment in which reward is expected (cue-reward consistencies are already known).

Since the seminal works on the observing response (e.g., Kelleher et al., 1962; Wilton & Clements, 1971), it has become obvious that animals like rats, pigeons, and starlings, as well as humans, are very sensitive to the opportunities to get information about an outcome. They are willing to pay a cost to know the outcome in advance, even if this information is irrelevant to future actions (Bennett et al., 2016; Eliaz & Schotter, 2010; Embrey et al., 2021; FitzGibbon et al., 2020; Rodriguez Cabrero et al., 2019). But one of the most compelling pieces of evidence that animals may value the consistency of a CS-US pairing more

than the US itself comes from the so-called suboptimal choice task (Kendall, 1974; Spetch et al., 1990; Zentall, 2016). The animal is given a choice between two options. Choosing one option has the immediate effect of revealing a cue, which *consistently* indicates whether the trial will end with food delivery (say, green cue) or no food delivery (say, red cue) with a 100% probability a few seconds later. The good news (green → food) occurs less often than the bad news (red → no food)—e.g., in 20% vs. 80% of the trials, respectively. Choosing the other option reveals one of two cues (yellow or blue cue) that are both *inconsistently* followed by food or no food with a 50% probability a few seconds later. Both cues have equal probability following choice. In short, the former option is suboptimal in terms of reward rate relative to the latter: 20% vs. 50%; 2.5 times less food received. However, pigeons and starlings (and rats under sometimes-different parametric values) strongly prefer the consistent, suboptimal option. This suboptimal preference is maintained despite enormous deficits in the reward amounts that can be collected relative to the optimal option (Fortes et al., 2016; Vasconcelos et al., 2015).

Several theories referring to the probabilistic contrast between what is expected and what occurs (Stagner & Zentall, 2010), delay reduction (McDevitt et al., 2016), and other phenomena (Zentall, 2019), have been proposed to account for those findings and others. Some results indicate that probabilistic contrast plays a role in this process (Zentall et al., 2019), but its effects are possibly too slow (they require more than 25 training sessions) to explain most experimental findings, where a suboptimal preference usually emerges after 10–15 training sessions. In addition, several results are incompatible with a prediction in terms of probabilistic contrast. For example, pigeons are indifferent in choice when both the suboptimal and optimal options are consistent, despite the presence of a contrast in the former (Smith & Zentall, 2016, Experiments 1 therein). Also, pigeons prefer a consistent optimal option to an inconsistent suboptimal option, despite the absence of a contrast in both (Smith & Zentall, 2016, Experiment 2 therein; see also Zentall et al., 2022). Hence, probabilistic contrast might essentially be a correlate rather than a cause of suboptimal preference. Alternatively, the theory based on delay reduction suggests that delay acts as a conditioned reinforcer in the suboptimal option when the rewarded cue (green) occurs, because it is good news that food is to be delivered soon relative to the non-rewarded cue (red). On the contrary, delay is not a conditioned reinforcer in the optimal option since both cues (blue and yellow) are associated with the same expectation of food delivery (McDevitt et al., 2016). This view accords well with evidence from temporal discounting experiments, showing that delays are typically aversive (Tobin & Logue, 1994; Mazur & Biondi, 2009). However, it does not explain why the probability of good vs. bad news (green vs. red) in the suboptimal option alters the strength of suboptimal choice (Zentall, 2016).

Kendall (1974) found that pigeons prefer a suboptimal option with consistent cues, but not a suboptimal option with inconsistent cues. It may seem surprising that, contrary to what Smith and Zentall (2016) found in their first experiment, this preference occurred despite an optimal option 100% consistent as well. However, it must be noted that in Kendall's study, food was delivered after a delay, which was for many sessions up to twice shorter in the suboptimal option than in the optimal one—until eventually reaching 15 s in both options. Preference is likely to have durably been influenced by this initial training. This may demonstrate the importance of short delays in suboptimal preference, a view that is compatible with consistency tracking: Shorter delays make a reward more likely, so animals may have evolved to perceive the predictive cue after a short delay as more consistent (Anselme, 2022a). Thus, the basic principle behind suboptimal choice might be that animals track the consistency of cue-food pairings (information) more than food (the reward), contrary to what the traditional rule of reward maximization suggests (Anselme, 2022a). And sensitivity to delays might be one component of this principle.

Interestingly, in humans, Liew et al. (2022) conducted three experiments in which they varied the cue-outcome delay, the valence of the outcome (reward vs. punishment), and the type of outcome (primary vs. monetary reinforcer). Then, they tested the ability of three models to predict non-instrumental information seeking i.e., an information that satisfies curiosity but has no consequences for future actions. The first two models were based on a prediction error mechanism, one suggesting that information seeking occurs because of the anticipation of a possible reward and the other that information seeking results from the anticipated learning signal. A third model instead posited that uncertainty is aversive and motivates organisms to resolve it as soon as possible. Their findings were clearly in favor of the uncertainty penalty model (third one) as, contrary to what the other two models predict, they found a monotonic increase of non-instrumental information seeking with the cue-outcome delay and no avoidance of that behavior in any

experimental condition. Liew et al. (2022) concluded that “when people seek noninstrumental information, they are seeking the resolution of uncertainty rather than savoring the anticipation of future outcomes” (p. 19). Said another way, under reward uncertainty, tracking cue-reward consistency motivates behavior and has a more direct influence on performance than the reward or the learned prediction. Although suboptimal choice is unrelated to the question of environmental exploration, I defend that, like in Liew et al.’s study, this task emphasizes the importance of tracking an information—the choice stimulus in this case—that leads to consistent cue-reward pairings (green → food or red → no food, in the example) and that is relevant to exploratory search. In the next section, I discuss information types that favor consistency tracking.

Conditioning, Higher-Order Conditioning, and Occasion Setting

A consistent cue-reward pairing always occurs in a context of other cues, which may help identify reward location or timing. One of these influences comes from higher-order conditioning, such as sensory preconditioning—where a neutral cue can acquire more incentive value than another, because it occurred in a chain of neutral stimuli that accidentally led to reward procurement. Specifically, the pairing of two neutral cues (A → X) followed by a pairing of one of them with a US such as food (X → US) produces a conditioned response (e.g., salivation) to X as well as to A, which was never paired with the food US (Holmes et al., 2022). The reverse procedure (X → US; A → X; A), called second-order conditioning, also generates a conditioned response to the non-paired cue. As pointed out by Honey and Dwyer (2022), higher-order conditioning “extends the ways in which Pavlovian conditioning can influence behavior to a broader range of real-world settings, where events with primary motivational significance (potential USs) are relatively rare” (p. 4-5). If sensory preconditioning may indeed sustain exploration, it is unlikely to explain what motivates exploration in the first place, however (e.g., why organisms will explore an environment containing stimuli never associated with any US through higher-order conditioning).

Any cue can potentially become a consistent CS, even if not directly “attached” to a US, like the color and shape of prey. For example, the smell of decay is a reliable predictor of a carcass nearby for a hyena, and squeaking is a reliable predictor of prey for a fox. In humans, a relevant example is that of a bird, the great honeyguide (*Indicator indicator*), which gets the attention of people from hunter-gatherer societies in sub-Saharan Africa. The bird guides those people to a bee colony, allowing them to collect honey and indirectly getting access to high-quality food i.e., bee eggs, larvae, and wax (Isack & Reyer, 1989). The sight of the great honeyguide (distal CS) is a likely guarantee for hunter-gatherers to find a bee colony (proximal CS) with honey reserves (food). The use of honeyguides reduces the search time for honey and considerably increases the chance of success (Spottiswoode et al., 2016; Wood et al., 2014).

Cues perhaps more directly informative about cue-reward consistency are feature cues, which act as occasion setters. An occasion setter (OS) is a stimulus whose presence or absence indicates whether another stimulus (CS) will be followed by reward or non-reward (Holland, 1992). In a feature-positive situation, the target CS (e.g., a tone) is only reinforced when preceded by a feature cue (e.g., flashing light). Conversely, a feature-negative situation means that the target CS is only reinforced in the absence of the feature cue. Owing to the disambiguating property of the feature cue, animals acquire an information about when to act or not to act, showing a conditioned response in the relevant context only (Bouton & Swartzentruber, 1986). Finding such cues is therefore of great help to learn the structure of an environment and identify reliable signals of good or bad news—and make appropriate decisions—during an exploratory bout. For example, the sight of a lion nearby causes the gazelle to flee, unless the lion is at the watering hole. Lions do not chase prey at a watering hole. Thus, the watering hole acts as a negative occasion setter that prevents the CS (lion) from eliciting fear in prey (gazelle) as it otherwise would.

OSs are sensitive to a number of Pavlovian phenomena (blocking, overshadowing, latent inhibition, etc.), and therefore resemble traditional CSs (Miller & Oberling, 1999). Fraser and Janak (2019) also showed that, like CSs, they can acquire incentive salience on their own right and increase that of otherwise undesirable CSs. However, OSs and CSs differ significantly in other respects. Contrary to CSs, OSs are very resistant to extinction and easily transfer from one CS-US association to another in a way that facilitates memory retrieval (Holland, 1992). Current evidence suggests that a stimulus is more likely to become an OS rather than a CS if the stimulus precedes the CS in time (no simultaneity, no overlap), the

time interval between the OS and the CS is long, the intertrial interval is long, the OS is less salient/intense than the CS, and the OS is presented in a sensory modality that differs from that of the CS (Fraser & Holland, 2019). A non-informative stimulus—i.e., a stimulus whose occurrence is random, or precedes a non-ambiguous CS, or does not resolve ambiguity in any way—cannot acquire the properties of an OS and plays no role in Pavlovian responding. OSs provide an information about the consistency of a CS-US pairing and should therefore be tracked and preferred when available.

Occasion setting may contribute to explain how suboptimal choice works. The choice stimulus—also called initial-link stimulus—in the suboptimal option might play the role of an OS because it indicates the option (“context”) that contains information, in the form of consistent cue-reward pairings. Of course, choosing suboptimally does not offer any guarantee to be rewarded but the color that will follow choice (green or red) will reliably inform the animal whether good or bad news can be expected. Given that pigeons seem to prefer information over no information (González & Blaisdell, 2021; see also Sears et al., 2022), the initial-link stimulus of the suboptimal option, in a sense, creates a context that disambiguates the situation, as opposed to the initial-link stimulus of the optimal option—which provides no information, and hence does not act as an OS. It could instead be argued that suboptimal choice occurs as a weird response to an “unnatural” design that never happens in a real-life setting (Fortes et al., 2016; Zentall, 2019). For sure, the design is unlikely to reflect any natural situation. However, the suboptimal response may be consequential to selective pressures that shaped the evolution of organisms and that somehow relate to the conditions presented in the suboptimal choice task. If selective pressures on consistency tracking exist, because the learning of cue-reward consistencies improves decision making, then suboptimal choice does not appear as weird as sometimes believed—and does not even appear suboptimal relative to this criterion (consistency rather than reward).

The importance of consistency tracking through OSs is not just about food but also about any other kind of reward, including “home” as a set of cues consistently predictive of shelter and more. Moving a beehive a few meters further out from its initial location disturbs the returning forager bees, because the landmarks surrounding the hive and necessary for its identification have changed—the first flights of a forager bee are exploratory and aim to learn those landmarks (Chittka, 2022). Such landmarks are not only required to identify one’s hive, as a set of consistent cues predictive of honey reserves, conspecifics, and shelter, but also to limit possible confusions with other hives close by.

Hoarding Behavior as an Example of Consistency Tracking

Consistency tracking might help explain a behavior observed in many animal species, including humans: hoarding behavior i.e., the propensity to store food outside of one’s body for later use, in locations hidden from the sight of potential competitors of the same or other species (e.g., Gerber et al., 2004; Pravosudov, 2006). This behavior may occur when food is abundant (in autumn and sometimes in spring) in order to accumulate reserves and avoid possible food shortage later (in winter and perhaps in anticipation of the breeding season). For example, parids hoard a lot of food before winter, and all the sooner when the latitude is high (Pravosudov, 2006). Also, contrary to non-hoarding animals like mice and rats, hoarding animals like Siberian and Syrian hamsters do not overeat when refed after a period of food deprivation; they hoard more (Bartness et al., 2011). The advantage associated with hoarding behavior is to remain leaner, that is, fast and agile in case of a predatory attack (Brodin, 2000).

However, what motivates this activity has hardly been questioned. In a study with scrub jays (*Aphelocoma coerulescens*), Clayton and Dickinson (1999) found that hoarding is only partly related to the feeding system. Hoarding is highly sensitive to spatial cues (Kamil & Balda, 1985) but also to non-spatial information capable of improving retrieval accuracy (Feenders & Smulders, 2011; LaDage et al., 2009). The function of hoarding might therefore be the reduction of variability in food availability to minimize the risk of starvation later (Sherry, 1985; Hitchcock & Houston, 1994): Hoarding considerably reduces the time and effort required to seek food when environmental unpredictability becomes too high, like in the middle of winter. The principle of consistency tracking is in line with such an explanation: The hoarder establishes consistent associations between specific signals (spatial cues or others) and some food items (rewards) in prevision of upcoming difficult times. To optimize the persistence of consistency in time, hoarders often scatter many small hoards (consistent associations)—a strategy known to limit the risk of pilferage (Male

& Smulders, 2007). Thus, hoarding behavior can reasonably be described as a behavior that establishes cue-reward consistencies through exploratory search. It is the result of selective pressures that favor good over bad foraging decisions in a hostile environment for the individuals of many species.

Incentive effort

Choosing an option that provides information about cue-reward consistency may be indicative of an organism's motivational target, but choice behavior is not exploratory as such. However, in a setting where choice is not the only possible behavior, reaching this motivational target relates to environmental exploration: cue-reward consistency cannot be approached, like a stimulus, but it can be sought. Nevertheless, appropriate stimuli have to be found, and two mechanisms should play a crucial role in this respect. One of them is called *incentive effort* here, and simply means that uncertainty makes effort self-motivating—the second mechanism is behavioral variability and will be discussed further. Attributing an incentive value to effort will strengthen and/or lengthen exploratory activity to compensate for the organism's lack of cognitive control in an uncertain environment.

Why is it necessary to refer to effort as self-motivating in a context of search under reward uncertainty? First, an inconsistent cue-reward pairing is less often chosen than its consistent counterpart in a simple free-choice task (bees: Anselme, 2018; pigs: de Jonge et al., 2008; macaques: Eisenreich et al., 2019; humans: Gneezy et al., 2006). This suggests that uncertainty is in no way a potent reward, if any. Second, incentive salience theory predicts that cue attraction derives from the salience of the paired reward, so that a reward-predictive cue should be “wanted” less when reward is unguaranteed or delayed rather than predictable and immediate (Anselme, 2021). Such a prediction is, in a sense, confirmed by probability and temporal discounting effects, where a cue receives fewer responses as it becomes less likely or more delayed (Green et al., 2014). Overall, those motivational and behavioral effects make it hard to explain how unguaranteed rewards can motivate exploratory search.

Instead, I defend that reward uncertainty stimulates seeking behavior because uncertainty is perceived as a challenge to be overcome or resolved. This perceived challenge is what makes effort self-motivating with positive consequences (perseverance-induced compensations), as shown below, which reinforce effort and make it adaptive under reward uncertainty. Inglis (2000) suggested that uncertainty reduction is reinforcing, although I rather interpret uncertainty reduction as a consequence of incentive effort in a challenging situation. Indeed, as shown further, uncertainty boosts responding in experimental procedures in which no uncertainty reduction is possible (e.g., Anselme et al., 2013; Bateson et al., 2021; Pravosudov & Grubb, 1997). In short, like consistency tracking, incentive effort contributes to uncertainty resolution but is reinforced (and adaptive) for a different reason: While consistency tracking maximizes good foraging decisions under reward uncertainty perceived as a source of insecurity, incentive effort results in greater perseverance with reward uncertainty perceived as a challenge. Incentive effort should statistically increase the chance of identifying potential CSs and OSs, and hence consistent cue-reward pairings. (Similarly, the probability of obtaining a 4 after rolling a die once is 1/6, but 4 can be obtained many times with certainty after multiple rolls of the die.) A computer model by Abrams (1991) predicts that optimal foraging effort should increase when the increase in food density is temporary, when the decrease in food density causes a risk of starvation, and when foraging aims to minimize mortality.

The Pavlovian autoshaping procedure may indicate that organisms perceive reward uncertainty as a challenge to meet, independently of the opportunity to reduce it or to obtain more reward. When a specific CS (lever presentation or key illumination) is inconsistent in predicting reward delivery, the response rates to that CS are typically increased compared to its consistent counterpart in animals from different species (rats: Anselme et al., 2013; quails: Crawford et al., 1985; pigeons: Gottlieb, 2004). This effect is robust; it has been replicated many times in rats and pigeons. The animals trained with an inconsistent CS do not receive more food, sometimes less, within a session than those trained with a consistent one. Also, the obtained food is free in each rewarded trial (no action is required) and the animal has no control in the task. If effort as a response to a challenge was not incentive on its own, why would animals deploy more non-rewarded effort under uncertainty in this task than expected in terms of incentive salience?

The most obvious explanation would be to say that the attribution of incentive salience is higher for an inconsistent CS than for a consistent one. Indeed, the release of dopamine in the brain regions

associated with incentive salience is proportional to the inconsistency of cue-food pairings in a Pavlovian preparation (Fiorillo et al., 2003; Hart et al., 2015). However, there are reasons to believe that the surplus of effort under reward uncertainty in a Pavlovian design is not a consequence of incentive salience (Anselme, 2021; 2022b). First, after showing higher lever pressing under uncertainty than certainty training, the incentive salience of the CS lever and the US reward was assessed separately in the two groups of drug-free rats. In both groups, the rats were similarly motivated to access their respective CS lever (inconsistent or consistent) in a conditioned reinforcement test, and they were similarly motivated to obtain their respective reward (uncertain or certain) despite an increasing action cost over the trials (Hellberg et al., 2018; Robinson et al., 2019). A higher investment following uncertainty training would have been expected in these two tasks, if incentive salience was responsible for the higher response rates shown in autoshaping. Second, if incentive salience controlled the higher performance under reward uncertainty in autoshaping, an inconsistent CS should be preferred to a consistent CS in a simple free-choice test involving the two stimuli. However, as already noted, this virtually never happens; from insects to mammals, the individuals typically avoid uncertainty when possible. Thus, despite an amplification of cue-triggered responses in autoshaping, reward uncertainty appears poorly appetitive in comparison with the certainty of receiving a reward. The additional effort shown toward an inconsistent CS in autoshaping is more likely to be the symptom of a general (default) response to uncertainty (Anselme & Güntürkün, 2019), which is perceived as a challenge to meet rather than a factor increasing incentive salience attribution to the CS.

When uncertainty is an intrinsic property of the environment, there is no opportunity to resolve it through learning. The only possible strategy to overcome its aversive effects maximally is to work hard. For example, we exposed pigeons to a semi-natural environment, consisting of a board perforated with 180 holes, in which grains could be hidden—a thin plastic tape covered the holes, with a cross cut to create an opening above each hole. We divided the 180-hole board in several areas, including one in which the holes were inconsistently baited (30 holes randomly baited out of 90) and one in which the holes were all consistently baited (30 holes baited out of 30). The pigeons ate the same amount of food in both configurations. But they spent more time and effort per visit in the inconsistent area, suggesting that this additional investment was not only required to find the food items when difficulty increased but also that the pigeons were willing to work more under reward uncertainty (Anselme et al., 2022). Overall, these results are not predicted by the matching law, since this view suggests that behavior should match reinforcement rate per hole. Investment should have been decreased (rather than increased) in the inconsistent area relative to the consistent area—of note, there was no penalty such as a time transit associated with area switching. As pointed out by Houston et al. (2021): “Although it is often stated that matching is optimal [...], this statement does not make sense without a specification of the nature of the decision processes and the environment” (p. 6). In an environment full of uncertainties, overmatching (responding more than expected, based on reward rate per hole only) is likely to be the optimal strategy to stay alive.

Contrafreeloading as an Example of Incentive Effort

Contrafreeloading denotes the fact that, under some circumstances, animals may prefer earned over free food. Forkman (1993) gave Mongolian gerbils (*Meriones unguiculatus*) a choice between a bowl containing 1000 seeds and a bowl containing only 200 seeds hidden in sand. The gerbils spent more time foraging and ate more seeds from the 200-seed than the 1000-seed bowl. In other words, they were willing to deploy some effort to seek uncertain food when a more profitable alternative was freely available—of note, they were relatively unattracted by a bowl of sand without seeds (Forkman, 1991). This phenomenon occurs only if the food items are not visible, forcing the individuals to estimate their number through exploratory search (Bean et al., 1999; Forkman, 1996). Why does this happen? Contrafreeloading has been identified in almost all species tested (Inglis et al., 1997) and seems unrelated to the pursuit of CSs or of larger amounts of food, so the incentive salience hypothesis is in trouble with explaining this higher-cost behavior for uncertain outcomes. Also, this activity is typically observed when animals are not hungry, suggesting that approaching food is not its main function (e.g., Inglis & Ferguson, 1986; Morgan, 1974; Neuringer, 1969).

One major factor that motivates contrafreeloading is the presence of reward uncertainty in the earned-food option (e.g., Anderson & Chamove, 1984; Forkman, 1993; Inglis et al., 1997). Once uncertainty has been resolved, the free-food option comes to be preferred (Havelka, 1956). The motivation behind contrafreeloading therefore seems to be uncertainty reduction in the environment (Inglis, 2000; Inglis et al., 1997). But what are the incentive stimuli? How can an animal be motivated to tackle uncertainty—as an aversive state—in the first place (rather than fleeing the situation)? And, as already noted with respect to autoshaping, how can uncertainty invigorate responding despite an absence of additional payoff? Of course, there are secondary benefits to contrafreeload, because the available reserves will be consumed and will have to be replaced by new hoards or good knowledge of the surroundings. But this represents an evolutionary (ultimate) explanation of *why*, not a mechanistic (proximate) explanation of *how*, secondary benefits can motivate such a behavior.

Contrafreeloading can easily be explained in reference to incentive effort: organisms exposed to reward uncertainty invest more time and effort to compensate for their lack of cognitive control in the situation. Their higher investment may lead to uncertainty reduction i.e., to a better cognitive control in the situation, allowing organisms “to improve and update their estimate of the profitability of an uncertain food source that may unpredictably become the optimal place to feed” (Inglis et al., 1997, p. 1187). But the motivation behind this higher investment is assumed to result from a perception of uncertainty as a challenge to meet. Contrafreeloading delays reward procurement and makes it unavailable without effort, suggesting that effort must be self-motivating to enable animals to go away from the free, immediate food.

Behavioral variability

Uncertainty results from the dynamical nature of the world, which prevents organisms from having omniscient knowledge of its causal structure—even if their memory capacities were unlimited. Organisms only experience covariances between events, which are rarely perfect. We saw that incentive effort is an effective way of seeking information to compensate for the lack of omniscient knowledge, as this process participates in increasing the chance of identifying profitable places within a reasonable time frame. Another strategy consists of producing variation in the response to a situation associated with unguaranteed outcome. Several findings indicate that reward expectation and response variation have a negative relationship (Blaisdell et al., 2016). A high expectation (high reward probability) decreases variation, reflecting exploitation of a known resource. By contrast, a low expectation (low reward probability) increases variation, reflecting exploration for other resources. Both in an instrumental and a Pavlovian procedure with pigeons, cues associated with reward probabilities ranging from 100% to 0.6% showed similar results: The greater the expectation of reward, the less variable was the response to the cue (Stahlman et al., 2010a,b). This phenomenon occurred with respect to peck location (spatial dimension) and the interpeck interval (temporal dimension).

Interestingly, in a semi-natural setting, such as an open field, similar patterns were observed with respect to the navigation of rats (Stahlman & Blaisdell, 2011). Sixteen cups containing sand were surmounted by small wood blocks with two different landmarks. In the trials with the high-food landmark, one cup was baited 100% of the time with a food item buried in sand, but this probability was only 20% in the trials with the low-food landmark. A measure of variability in total number of cups inspected before the rats searched in the cued cup showed that they explored more in the low-food than in the high-food trials.

At first glance, introducing variability in responding may seem equivalent to putting more effort in the task, since varying responses appears more costly than always using the same response. But variability and effort are not the same thing. Pressing a lever more under reward uncertainty does not require any variability in the responses, and vice versa. Also, variability is inversely proportional to probability, p , while effort is proportional to uncertainty, u , expressed as $p(1 - p)$ in probabilistic terms. These distinct properties make sense in a natural setting. Effort is worth being produced when reward is locally uncertain but globally expected. In this case, organisms can eventually obtain the sought resource with enough investment in the task (e.g., Anselme et al., 2022). By contrast, if response variability increases the chance of encountering new opportunities, varying the location and the time of a food-seeking response may help organisms detect alternative food sources instead of persisting in a dead-end strategy. New opportunities are beneficial to survival, making behavioral variability both reinforced and adaptive. Wittek et al. (2022) let pigeons forage

on a board of 180 holes randomly baited with 60 grains (33.3% holes with food) or on a board of 60 holes baited with 60 grains (100% holes with food). We observed within the same pigeons that the board with inconsistent pairings promoted a more uniform exploration of the surface available (with fewer pecks and fewer revisits per hole) and, perhaps more importantly here, a significant propensity to inspect a distant rather than adjacent hole after a peck. The traditional explanation of these effects in behavioral ecology consists of suggesting that the next expected reward per hole on the 180-hole board being lower than on the 60-hole board, less time and energy will be invested in searching (Charnov, 1976). In psychology, the matching law would predict something similar: Since behavior is proportional to reward rate, a poorer environment per hole should generate less investment than a richer one (Herrnstein, 1961)—although the absence of choice between both boards makes any interpretation in terms of matching impossible. However, if one board was less attractive than the other, how to explain that the pigeons were able to extract a similar number of food items after similar duration in both environments? The greater avoidance of adjacent holes, as well as the lower number of pecks and revisits per hole, on the 180-hole board, might indicate a greater variability in the pigeons' responses. In other words, a lower probability of food per hole optimized and spatially extended foraging activity.

Relocation as a Form of Behavioral Variability

Behavioral variability can also be discussed in a context that differs from that of the studies reported above, and the neurobehavioral mechanisms causing variability in this context may be different. If random search for food in a restricted patch is unsuccessful, should organisms extend their search to adjacent areas or should they move further away—when the environment is vast enough? The marginal value theorem (Charnov, 1976) suggests that organisms should move in case the cost of staying becomes higher than that of leaving. However, using landmarks indicating the next profitable patch is not always possible. Sharks navigating in troubled water or albatrosses flying over the ocean are in this situation, as well as animals with low memory and planning capacities such as mussels and snails. Many studies report that animals unable to find food locally tend to stop and move in a relatively straight-line direction, selected randomly, to relocate and start a new local search bout. Technically, local search is referred to as Brownian motion, which means that small step (or walk) lengths are displayed in random directions. This strategy maintains the individual close to its initial coordinates. By contrast, global search requires Lévy motion, where large step (or walk) lengths occasionally occur and relocate the individual far from its initial coordinates.

Although Brownian and Lévy motions are properties of very simple physical systems (such as the movements of dust particles in a fluid), the navigation of foraging animals resembles those patterns in many respects. There is a debate about the existence of Lévy walks in foraging organisms (e.g., Reynolds, 2015; Sims, 2015; Ziburdaev et al., 2015) and this paper is not the place to discuss this question. However, whether organisms perform Lévy or Lévy-like walks, their ability to relocate in case of local food shortage has been documented in many species, from protists to human hunter-gatherers (e.g., albatrosses: Humphries et al., 2012; snails: Kölzsch et al., 2015; humans: Raichlen et al., 2014; monkeys: Ramos-Fernandez et al., 2004; amoebae: Van Haastert & Bosgraaf, 2009).

Is this strategy adaptive? Here also, there is a debate because the exact mechanisms responsible for Lévy motions are unclear. However, Lévy motions were shown to optimize search efficiency in very simple environment (Viswanathan et al., 1999) and more recent statistical methods applied to large data sets involving long movement paths, support the hypothesis that Lévy motions have an adaptive value in two species of albatrosses (Humphries et al., 2012, 2013). de Jager et al. (2014) showed that Brownian motion is not a default movement pattern but rather emerges from frequent ecological interactions between organisms: A food-rich environment will contain more foragers, causing more encounters, and hence more “collisions” that will mechanically result in Brownian motions. The adoption of a Lévy or Brownian motion would therefore be unrelated to a change in food density *per se*, as traditionally believed. However, Abe (2020) notably showed that Lévy walks are involved in a switch between exploration and exploitation depending on the sensory information received, providing an adaptive advantage to environmental conditions. The sensory information may be received from the environment but also have an internal origin.

In computer simulations, we found that the motivation to seek food plays a role in the occurrence of Lévy walks (Anselme et al., 2018). Our foraging agents—one per simulation, no collisions possible—

produced Lévy walks more often under an unpredictable than a predictable food access, and more often when their motivation to seek food was high rather than low in the unpredictable condition. Our results indicate that Lévy walks may confer some advantages for survival in an unpredictable environment, if they appear in foragers with a high motivation to seek food. Using the same computer program (without an analysis of motions) in other studies, we confirmed that food seeking was more spatially extended (variable) and, as a result, the individual agents ate more food items, had higher energy reserves, and survived longer under a fourfold lower food density in a situation of semi-destructive search i.e., one food item consumed was replaced but in a different, random location (Anselme et al., 2017; Anselme & Güntürkün, 2019).

A need to know or the perception of challenges?

Active information seeking is often assumed to depend on a so-called “need to know.” However, this formulation is problematic in several respects. I would like to show that the three principles described in the present paper make active information seeking possible in the absence of a need to know.

First, a need to know would imply that any kind of information is a reward to be obtained. However, organisms are surrounded by myriads of informational contents, and most of them are virtually ignored. If information was rewarding by itself, we should all spend considerable amounts of time reading the Civil Code or the telephone book. People do not usually see the rules contained in the Highway Code or their application as rewards. Therefore, information can motivate behavior in a restricted manner only. In this paper, I defend that the major type of information actively sought by organisms, including humans, is the consistency of cue-reward (or action-reward) pairings. Most other informational contents may passively emerge from the consequences of behavior (see Borgstede, 2021). Second, the concept of need to know is rather enigmatic, and reminds the no less enigmatic concept of behavioral need. Let us consider the case of behavior. It is common to see the various pathologies (behavioral stereotypes, gastric ulcers, and so forth) developed by captive animals in zoos as the consequence of their inability to express their “behavioral needs” (Hughes & Duncan, 1988)—such as running for cheetahs or digging for armadillos. After all, they receive enough food and care from their keeper; they are not hungry and do not suffer from any other physiological deprivation. However, the concept of behavioral need has been criticized, notably because it relies on the assumption that behavior has internal causes only rather than being a combination of internal and external factors (Jensen & Toates, 1993). Also, as reported by these authors, there is no reason to consider a motion pattern as a need, even in case it is self-rewarding—e.g., chasing a ball may be fun for a dog but do we have to postulate a need to chase balls (or anything else) in dogs? The same problems occur with respect to the need to know: There is no internal sensitivity to information without external cues and rewards, and the self-rewarding dimension of information seeking does not require any need. Also, more generally, the fact that we can invent as many needs as we want to explain behavior, makes this notion useless. If any “need” had to be invoked (like when you infer a need for water from the fact that life is not possible in the absence of water), it would be a need to meet challenges instead of a need to know. In this sense, incentive effort and response variability are a source of good health (like water drinking is) because they are the exploratory strategies used by animals to meet challenges in their environment.

The inability for most captive animals to express incentive effort and response variability deteriorates their general health because they fail to receive positive feedback from these activities as responses to environmental challenges that never happen. Why are these challenges salutary for health (rather than situations that are worth avoiding)? Because (most) organisms evolved in challenging environments, and both their physiology and psychology were shaped by natural selection to do this job. An under-stimulated captive animal has its well-being at stake, just as a severely thirsty animal without access to water is at risk. Needs are high-level descriptions that we use to easily understand why captive animals develop pathologies and why thirsty animals show polydipsia symptoms, rather than basic internal states triggering specific behaviors. There is no need to know to be satisfied but only positive feedback normally received in challenging situations, which favor the expression of incentive effort and response variability. Accordingly, improvements in well-being through environmental enrichment were shown in captive mice and pigs, which had more opportunities to be exposed to various stimuli, explore, and achieve their goals than individuals under standard housing (Beattie et al., 2000; Marashi et al., 2003). Improvements in well-being and emotional/cognitive skills have also been reported after episodes of video

gaming in humans (Pallavicini et al., 2018). By contrast, playing a game without uncertainty about the path to follow or the outcome is dull because of the absence of challenges in the task (Anselme & Robinson, 2013; Costikyan, 2013).

Conclusion

From a motivational perspective, exploratory search must be distinguished from the well-studied resource exploitation strategies. During exploration, the individual's motivation does not consist of approaching cues and rewards but aims at resolving uncertainty through information seeking. In this paper, I showed that three general psychological principles may underpin this process. Whatever the exact neuronal implementation of consistency tracking, incentive effort, and response variability, they point to what motivates exploration and how organisms can reach their target in an uncertain or unknown environment. Uncertainty maximizes their effects, not because there is something appetizing with unguaranteed outcomes, but because uncertainty creates a challenging situation that organisms may try to overcome or resolve. Organisms evolved to meet environmental challenges and their good health depends on their opportunity to be exposed to survival-related problems regularly. Future studies on information seeking might help characterize exploratory search in more details and explain how it works in concert with resource exploitation.

Funding. This work was supported by the Deutsche Forschungsgemeinschaft (DFG), Germany, through An1067/3-1.

Conflict of interest. The author declares no conflict of interest.

References

- Abe, M.S. (2020). Functional advantages of Lévy walks emerging near a critical point. *Proceedings of the National Academy of Sciences*, *117*, 24336-24344.
- Abrams, P.A. (1991). Life history and the relationship between food availability and foraging effort. *Ecology*, *72*, 1242-1252.
- Anderson, J. R. & Chamove, A. S. (1984). Allowing captive primates to forage. In: *Standards in Laboratory Animals Management* (Part 2, pp. 253–256). Potters Bar, UK: Universities Federation of Animal Welfare.
- Anselme, P. (2018). Uncertainty processing in bees exposed to free choices: Lessons from vertebrates. *Psychonomic Bulletin and Review*, *25*, 2024–2036.
- Anselme, P. (2021). Effort-motivated behavior resolves paradoxes in appetitive conditioning. *Behavioural Processes*, *193*, 104525.
- Anselme, P. (2022a). The optimality of “suboptimal” choice: A psycho-evolutionary perspective. In: M. Krause, K.L. Hollis, M.R. Papini (eds.), *Evolution of Learning and Memory Mechanisms* (chapter 11, pp. 193-209). Cambridge: Cambridge University Press.
- Anselme, P. (2022b). Reward uncertainty and the aversion-attraction dilemma. In: A.J. Elliot (ed.), *Advances in Motivation Science* (vol. 9, pp. 1-34). New York: Elsevier.
- Anselme, P. & Robinson, M. J. F. (2013). What motivates gambling behavior: Insight into dopamine's role. *Frontiers in Behavioral Neuroscience* *7*, 182.
- Anselme, P. & Güntürkün, O. (2019). How foraging works: uncertainty magnifies food seeking motivation. *Behavioral and Brain Sciences*, *42*: e35, 1-59.
- Anselme, P., Robinson, M.J.F. & Berridge, K.C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behavioural Brain Research*, *238*, 53-61.
- Anselme, P., Otto, T. & Güntürkün, O. (2017). How unpredictable access to food increases the body fat of small passerines: A mechanistic approach. *Behavioural Processes*, *144*, 33–45.
- Anselme, P., Otto, T. & Güntürkün, O. (2018). Foraging motivation favors the occurrence of Lévy walks. *Behavioural Processes*, *147*, 48-60.
- Anselme, P., Wittek, N., Oeksuez, F., & Güntürkün, O. (2022). Overmatching under food uncertainty in foraging pigeons. *Behavioural Processes*, *201*, 104728.
- Bartness, T. J., Keen-Rhinehart, E., Dailey, M. J. & Teubner, B. J. (2011). Neural and hormonal control of food hoarding. *American Journal of Physiology* *301*, R641–R655.

- Bartumeus, F., Campos, D., Ryu, W. S., Lloret-Cabot, R., Méndez, V. & Catalan, J. (2016). Foraging success under uncertainty: Search tradeoffs and optimal space use. *Ecology Letters*, *19*, 1299–313.
- Bateson, M., Andrews, C., Dunn, J., Egger, C.B.C.M., Gray, F., Mchugh, M. & Nettle, D. (2021). Food insecurity increases energetic efficiency, not food consumption: an exploratory study in European starlings. *Peer J*, *9*, e11541.
- Baum, W.M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, *22*, 231–242.
- Bean, D., Mason, G.J. & Bateson, M. (1999). Contrafreeloading in starlings: testing the information hypothesis. *Behaviour*, *136*, 1267-1282.
- Beattie, V. E., O’Connell, N. E., Kilpatrick, D. J. & Moss, B. W. (2000). Influence of environmental enrichment on welfare-related behavioural and physiological parameters in growing pigs. *Animal Science*, *70*, 443- 450.
- Beeler, J.A. & Mourra, D. (2018). To do or not to do: dopamine, affordability and the economics of opportunity. *Frontiers in Integrative Neuroscience*, *12*, 6.
- Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS Computational Biology*, *12*, e1005020.
- Berlyne, D. E. (1960). Conflict, arousal, and curiosity. McGraw-Hill Book Company.
- Berridge, K.C. (2007). The debate over dopamine’s role in reward: the case for incentive salience. *Psychopharmacology*, *191*, 391–431.
- Blaisdell, A.P., Stolyarova, A., & Stahlman, W.D. (2016). The law of expect or a modified law of effect? Outcome expectation and variation in learned behavior. *Conducta*, *4*, 61-90.
- Bond, A.B. (1983). Visual search and selection of natural stimuli in the pigeon: the attentional threshold hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 292–306.
- Borgstede, M. (2021). Why do individuals seek information? A selectionist perspective. *Frontiers in Psychology*, *12*, 684544.
- Bouton, M.E. & Swartzentruber, D. (1986). Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes* *12*, 333-350.
- Brodin, A. (2000). Why do hoarding birds gain fat in winter in the wrong way? Suggestions from a dynamic model. *Behavioral Ecology*, *11*, 27-39.
- Charnov, E.L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*, 129-136.
- Chatterjee, A., Bais, D., Brockmann, A., & Ramesh, D. (2021). Search behavior of individual foragers involves neurotransmitter systems characteristic for social scouting. *Frontiers in Insect Science*, *1*, 664978.
- Chittka, L. (2022). *The mind of a bee*. Princeton: Princeton University Press.
- Cook, C. N., Mosquero, T., Brent, C. S., Ozturk, C., Gadau, J., Pinter-Wollman, N., & Smith, B. H. (2019). Individual differences in learning and biogenic amine levels influence the behavioural division between foraging honeybee scouts and recruits. *Journal of Animal Ecology*, *88*, 236-246.
- Clayton, N.S. & Dickinson, A. (1999). Motivational control of caching behaviour in the scrub jay, *Aphelocoma coerulescens*. *Animal Behaviour*, *57*, 435-444.
- Costikyan, G. (2013). *Uncertainty in Games*. Cambridge, MA: MIT Press.
- Crawford, L.L., Steirn, J.N., & Pavlik, W.B. (1985). Within- and between-subjects partial reinforcement effects with an autoshaped response using Japanese quail (*Coturnix coturnix japonica*). *Animal Learning and Behavior*, *13*, 85–92.
- de Jager, M., Bartumeus, F., Kölzsch, A., Weissing, F.J., Hengeveld, G.M., Nolet, B.A., et al. (2014). How superdiffusion gets arrested: ecological encounters explain shift from Lévy to Brownian movement. *Proceedings of the Royal Society, B* *281*, 20132605.
- de Jonge, F. H., Ooms, M., Kuurman, W. W., Maes, J. H. R., & Spruijt, B. M. (2008). Are pigs sensitive to variability in food rewards? *Applied Animal Behaviour Science*, *114*, 93–104.
- DiFeliceantonio, A., Mabrouk, O.S., Kennedy, R.T., & Berridge, K.C. (2012). Enkephalin surges in dorsal neostriatum as a signal to eat. *Current Biology*, *22*, 1918-1924.
- Eisenreich, B. R., Hayden, B. Y., & Zimmermann, J. (2019). Macaques are risk-averse in a freely moving foraging task. *Scientific Reports*, *9*, 15091.
- Eliasz, K. & Schotter, A. (2010). Paying for confidence: an experimental study of the demand for non-instrumental information. *Games and Economic Behavior*, *70*, 304–324.
- Embrey, J.R., Liew, S.X., Ghai, I., & Newell, B.R. (2021). The effect of uncertainty and outcome probability on non-instrumental information seeking. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*, 1257-1263.
- Feenders, G. & Smulders, T.V. (2011). Magpies can use local cues to retrieve their food caches. *Animal Cognition*, *14*, 235-243.

- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*, 1898–1902.
- FitzGibbon, L., Lau, J.K.L. & Murayama, K. (2020). The seductive lure of curiosity: information as a motivationally salient reward. *Current Opinion in Behavioral Sciences*, *35*, 21–27.
- Forkman, B. (1991). Some problems with current patch choice theory: a study on the Mongolian gerbil. *Behaviour*, *117*, 243–254.
- Forkman B. (1993). The effect of uncertainty on the food intake of the Mongolian gerbil. *Behaviour*, *124*, 197–206.
- Forkman, B. (1996). The foraging behaviour of Mongolian gerbils: a behavioural need or a need to know? *Behaviour*, *133*, 129–143.
- Fortes, I., Vasconcelos, M., & Machado, A. (2016). Testing the boundaries of “paradoxical” predictions: pigeons do disregard bad news. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 336–346.
- Fraser, K. M., & Holland, P. C. (2019). Occasion setting. *Behavioral Neuroscience*, *133*, 145–175.
- Fraser, K.M. & Janak, P.H. (2019). Occasion setters attain incentive motivational value: implications for contextual influences on reward-seeking. *Learning and Memory*, *26*, 291–298.
- Gerber, L.R., Reichman, O.J., & Roughgarden, J. (2004). Food hoarding: Future value in optimal foraging decisions. *Ecological Modelling*, *175*, 77–85.
- Gneezy U., List J.A. & Wu G. (2006). The uncertainty effect: when a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics*, *121*, 1283–1309.
- González, V.V., & Blaisdell, A.P. (2021). The role of inhibition in the suboptimal choice task. *Journal of Experimental Psychology: Animal Learning and Cognition*, *47*, 429–444.
- Gottlieb, D.A. (2004). Acquisition with partial and continuous reinforcement in pigeon autoshaping. *Learning and Behavior*, *32*, 321–334.
- Green, L., Myerson, J., & Vanderveldt, A. (2014). Delay and probability discounting. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley Blackwell handbook of operant and classical conditioning* (pp. 307–337). Wiley Blackwell.
- Hart A. S., Clark J. J. & Phillips P. E. M. (2015). Dynamic shaping of dopamine signals during probabilistic Pavlovian conditioning. *Neurobiology of Learning and Memory*, *117*, 84–92.
- Havelka J. (1956). Problem-seeking behaviour in rats. *Canadian Journal of Psychology*, *10*, 91–97.
- Hebb, D.O. (1955). Drives and the C. N. S. (conceptual nervous system). *Psychological Review*, *62*, 243–254.
- Hellberg, S.N., Levit, J.D., & Robinson, M.J.F. (2018). Under the influence: effects of adolescent ethanol exposure and anxiety on motivation for uncertain gambling-like cues in male and female rats. *Behavioural Brain Research*, *337*, 17–33.
- Herrnstein, R.J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267–272.
- Hitchcock, C.L. & Houston, A.I. (1994). The value of a hoard: not just energy. *Behavioral Ecology*, *5*, 202–205.
- Holland, P.C. (1992). Occasion setting in Pavlovian conditioning. In: D.L. Medin (eds.), *The psychology of learning and motivation* (vol. 28, pp. 69–125). San Diego, CA: Academic Press.
- Holmes, N. M., Wong, F. S., Bouchekioua, Y., & Westbrook, R. F. (2022). Not “either-or” but “which-when”: A review of the evidence for integration in sensory preconditioning. *Neuroscience and Biobehavioral Reviews*, *132*, 1197–1204.
- Honey, R. C., & Dwyer, D. M. (2022). Higher-order conditioning: A critical review and computational model. *Psychological Review*, *129*, 1338–1357.
- Houston, A.I., Trimmer, P.C., & McNamara, J.M. (2021). Matching behaviours and rewards. *Trends in Cognitive Sciences*, *25*, 403–415.
- Huang, J., Zhang, Z., Feng, W., Zhao, Y., Aldanondo, A.,... Su, S. (2022). Food wanting is mediated by transient activation of dopaminergic signaling in the honey bee brain. *Science*, *376*, 508–512.
- Hughes, B.O. & Duncan, I.J.H. (1988). The notion of ethological “need”, models of motivation and animal welfare. *Animal Behaviour*, *36*, 1696–1707.
- Humphries, N.E., Weimerskirch, H., Queiroz, N., Southall, E.J., & Sims, D.W. (2012). Foraging success of biological Lévy flights recorded in situ. *Proceedings of the National Academy of Science*, *109*, 7169–7174.
- Humphries, N.E., Weimerskirch, H., & Sims, D.W. (2013). A new approach for objective identification of turns and steps in organism movement data relevant to random walk modelling. *Methods in Ecology and Evolution*, *4*, 930–938.
- Inglis, I. (2000). The central role of uncertainty reduction in determining behaviour. *Behaviour*, *137*, 1567–1599.
- Inglis, I. R. & Ferguson, N. J. K. (1986). Starlings search for food rather than eat freely available, identical food. *Animal Behaviour*, *34*, 614–617.
- Inglis, I.R., Forkman, B., & Lazarus, J. (1997). Free food or earned food? A review and fuzzy model of contrafreeloading. *Animal Behaviour*, *53*, 1171–1191.

- Isack, H. A. & Reyer, H.-U. (1989). Honeyguides and honey gatherers: interspecific communication in a symbiotic relationship. *Science*, *243*, 1343–1346.
- Jander, R. (1997). Macroevolution of a fixed action pattern for learning: the exploratory flights of bees and wasps. In G. Greenberg & E. Tobach (eds.), *Comparative psychology of invertebrates: The field and laboratory study of insect behavior* (pp. 79–99). New York, NY: Garland.
- Jensen, P. & Toates, F.M. (1993). Who needs “behavioural needs”? Motivational aspects of the needs of animals. *Applied Animal Behaviour Science*, *37*, 161–181.
- Kamil, A.C., & Balda, R.P. (1985). Cache recovery and spatial memory in Clark’s nutcrackers (*Nucifraga columbiana*). *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 95–111.
- Kelleher, R.T., Riddle, W.C., & Cook, L. (1962). Observing responses in pigeons. *Journal of the Experimental Analysis of Behavior*, *5*, 3–13.
- Keller, A.M., Taylor, H.A., & Brunyé, T.T. (2020). Uncertainty promotes information-seeking actions, but what information? *Cognitive Research: Principles and Implications*, *5*, 42.
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, *21*, 463–473
- Kölzsch, A., Alzate, A., Bartumeus, F., de Jager, M., Weerman, E.J., Hengeveld, G.M., Naguib, M., Nolet, B.A., & van de Koppel, J. (2015). Experimental evidence for inherent Lévy search behaviour in foraging animals. *Proceedings of the Royal Society, B* *282*, 20150424.
- LaDage, L.D., Roth II, T.C., Fox, R.A. & Pravosudov, V.V. (2009). Flexible cue use in food-caching birds. *Animal Cognition*, *12*, 419–426.
- Liang, Z. S., Nguyen, T., Mattila, H. R., Rodriguez-Zas, S. L., Seeley, T.D., & Robinson, G. E. (2012). Molecular determinants of scouting behavior in honey bees. *Science*, *335*, 122–125.
- Liew, S.X., Embrey, J.R., Navarro, D.J., & Newell, B.R. (2022). Comparing anticipation and uncertainty-penalty accounts of non-instrumental information seeking. *Decision*. Advance online publication.
- Lindauer, M. (1952). Ein Beitrag zur Frage der Arbeitsteilung im Bienenstaat. *Zeitschrift für vergleichende Physiologie*, *34*, 299–345.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*, 75–98.
- MacArthur, R.H. & Pianka, E.R. (1966). On the optimal use of a patchy environment. *American Naturalist*, *100*, 603–609.
- Male, L.H., & Smulders, T.V. (2007). Hyperdispersed cache distributions reduce pilferage: a field study. *Animal Behaviour*, *73*, 717–726.
- Marashi, V., Barnekow, A., Ossendorf, E. & Sachser, N. (2003). Effects of different forms of environmental enrichment on behavioral, endocrinological, and immunological parameters in male mice. *Hormones and Behavior*, *43*, 281–292.
- Mazur, J. E., & Biondi, D. R. (2009). Delay-amount tradeoffs in choices by pigeons and rats: Hyperbolic versus exponential discounting. *Journal of the Experimental Analysis of Behavior*, *91*, 197–211.
- McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, *105*, 23–40.
- McNickle, G.G., St. Clair, C.C., Cahill, J.F. (2009). Focusing the metaphor: plant root foraging behaviour. *Trends in Ecology and Evolution*, *24*, 419–426.
- Miller, R. R., & Oberling, P. (1998). Analogies between occasion setting and Pavlovian conditioning. In N. A. Schmajuk & P. C. Holland (eds.), *Occasion setting: Associative learning and cognition in animals* (p. 3–35). American Psychological Association.
- Morgan, M. J. (1974). Do rats like to work for their food? *Learning and Motivation*, *5*, 352–368.
- Neuringer, A. J. (1969). Animals respond for food in the presence of free food. *Science*, *16*, 399–401.
- Palminter, R.D. (2008). Dopamine signaling in the dorsal striatum is essential for motivated behaviors: Lessons from dopamine-deficient mice. *Annals of the New York Academy of Sciences*, *1129*, 35–46.
- Pallavicini, F., Ferrari, A., & Mantovani, F. (2018). Video games for well-being: A systematic review on the application of computer games for cognitive and emotional training in the adult population. *Frontiers in Psychology*, *9*, 2127.
- Pisula, W. (2009). *Curiosity and information seeking in animal and human behavior*. Florida: BrownWalker Press.
- Pravosudov, V. V. (2006). On seasonality in food-storing behaviour in parids: Do we know the whole story? *Animal Behaviour*, *71*, 1455–60.
- Pravosudov, V. V. & Grubb Jr., T. C. (1997). Management of fat reserves and food caches in tufted titmice (*Parus bicolor*) in relation to unpredictable food supply. *Behavioral Ecology*, *8*, 332–339.
- Pyke, G.H. (1984). Optimal foraging theory: a critical review. *Annual Review of Ecology and Systematics*, *15*, 523–575.

- Rager, C. L., Bhat, U., Bénichou, O., & Redner, S. (2018). The advantage of foraging myopically. *Journal of Statistical Mechanics: Theory and Experiment*, 7, 073501.
- 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 Science*, 111, 728–733.
- Ramos-Fernandez, G., Mateos, J.L., Miramontes, O., Cocho, G., Larralde, H., & Ayala-Orozco, B. (2004). Lévy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). *Behavioral Ecology and Sociobiology*, 55, 223–230.
- Reynolds, A. (2015). Liberating Lévy walk research from the shackles of optimal foraging. *Physics of Life Reviews*, 14, 59–83.
- Robinson, M.J.F., & Berridge, K.C. (2013). Instant transformation of learned repulsion into motivational ‘wanting’. *Current Biology*, 23, 282–289.
- Robinson, M.J.F., Clibanoff, C., Freeland, C.M., Knes, A.S., Cote, J.R., & Russell, T.I. (2019). Distinguishing between predictive and incentive value of uncertain gambling-like cues in a Pavlovian autoshaping task. *Behavioural Brain Research*, 371, 111971.
- Rodriguez Cabrero, J.A.M.R., Zhu, J.-Q., & Ludvig, E.A. (2019). Costly curiosity: people pay a price to resolve an uncertain gamble early. *Behavioural Processes*, 160, 20–25.
- Salamone, J.D. & Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, 137, 3–25.
- Sansone, C. & Harackiewicz, J.M. (2000). Intrinsic and extrinsic motivation: The search for optimal motivation and performance. Academic Press.
- Saunders B. T. & Robinson T. E. (2012). The role of dopamine in the accumbens core in the expression of Pavlovian-conditioned responses. *European Journal of Neuroscience*, 36, 2521–32.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80, 1–27.
- Sears, B., Dunn, R.M., Pisklak, J.M., Spetch, M.L., & McDevitt, M.A. (2022). Good news is better than bad news, but bad news is not worse than no news. *Learning and Behavior*, 50, 482–493.
- Seeley, T.D. (1983). Division of labor between scouts and recruits in honeybee foraging. *Behavioral Ecology and Sociobiology*, 12, 253–259.
- Sherry, D.F. (1985). Food storage by birds and mammals. *Advances in the Study of Behavior*, 15, 153–188.
- Sims, D.W. (2015). Intrinsic Lévy behaviour in organisms –searching for a mechanism: comment on Liberating Lévy walk research from the shackles of optimal foraging by A.M. Reynolds. *Physics of Life Reviews*, 14, 111–114.
- Smith, A.P., & Zentall, T.R. (2016). Suboptimal choice in pigeons: choice is primarily based on the value of the conditioned reinforcers rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 212–220.
- Spetch, M. L., Belke, T. W., Barnett, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior*, 53, 219–234.
- Spottiswoode, C. N., Begg, K. S., Begg, C. M. (2016). Reciprocal signaling in honeyguide-human mutualism. *Science*, 353, 387–389.
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin and Reviews*, 17, 412–416.
- Stahlman, W. D., Roberts, S., & Blaisdell, A. P. (2010a). Effect of reward probability on spatial and temporal variation. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 77–91.
- Stahlman, W. D., Young, M. E., & Blaisdell, A. P. (2010b). Response variability in pigeons in a Pavlovian task. *Learning & Behavior*, 38, 111–118.
- Stahlman, W.D., & Blaisdell, A.P. (2011). Reward probability and the variability of foraging behavior in rats. *International Journal of Comparative Psychology*, 24, 168–176.
- Stephens, D.W., & Krebs, J.R. (1986). Foraging Theory. Princeton: Princeton University Press.
- Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin and Review*, 1, 405–420.
- Tobin, H., & Logue, A. W. (1994). Self-control across species (*Columba livia*, *Homo sapiens*, and *Rattus norvegicus*). *Journal of Comparative Psychology*, 108, 126–133.
- Van Haastert, P.J.M., & Bosgraaf, L. (2009). Food searching strategy of amoeboid cells by starvation induced run length extension. *PLoS One*, 4, e6814.
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific Reports*, 5, 13874.

- 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*, 911–914.
- Whishaw, I.Q., Gharbawie, O.A., Clark, B.J., & Lehmann, H. (2006). The exploratory behavior of rats in an open environment optimizes security. *Behavioural Brain Research*, *171*, 230–239.
- Wilton, R.N., & Clements, R.O. (1971). Observing responses and informative stimuli. *Journal of the Experimental Analysis of Behavior*, *15*, 199–204.
- Wittek, N., Oeksuez, F., Güntürkün, O., & Anselme, P. (2022). More opportunities to peck for identical food availability increases foraging efficiency in pigeons. *Behaviour*, *159*, 1201–1224.
- Wood, B. M., Pontzer, H., Raichlen, D. A., & Marlowe, F. W. (2014). Mutualism and manipulation in Hadza–honeyguide interactions. *Evolution and Human Behavior*, *35*, 540–546.
- Yip, V., Beekman, M., & Latty, T. (2014). Foraging strategies of the acellular slime moulds *Didymium iridis* and *Didymium bahiense*. *Fungal Ecology*, *11*, 29–36.
- Zaburdaev, V., Denisov, S., & Klafter, J. (2015). Lévy walks. *Reviews of Modern Physics*, *87*, 483–530.
- Zentall, T. R. (2016). Resolving the paradox of suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 1–14.
- Zentall, T.R. (2019). What suboptimal choice tells us about the control of behavior. *Comparative Cognition & Behavior Reviews*, *14*, 1–17.
- Zentall, T. R., Andrews, D. M., & Case, J. P. (2019). Contrast between what is expected and what occurs increases pigeons' suboptimal choice. *Animal Cognition*, *22*, 81–87.
- Zentall, T.R., Peng, D.N., & Mueller, P.M. (2022). Pigeon's choice depends primarily on the value of the signal for the outcome rather than its frequency or contrast. *Journal of Experimental Psychology: Animal Learning and Cognition*, *48*, 135–144.