**When is it time to move on? Patch-leaving behavior during probabilistic foraging in humans and gerbils**

**Lasse Güldener\*1, Parthiban Saravanakumar2, Max F. K. Happel 2,3,5 , Frank W. Ohl2,3,4, Maike Vollmer.2,6,7, & Stefan Pollmann† 1,3**

1 Department of Experimental Psychology, Institute of Psychology, Otto-von-Guericke-University, 39106 Magdeburg, Germany

2 Departments of Systems Physiology of Learning, Leibniz-Institute for Neurobiology, 39118 Magdeburg, Germany

3 Center for Behavioral Brain Sciences, 39106 Magdeburg, Germany

4 Institute of Biology, Otto-von-Guericke-University, 39106 Magdeburg,

5 Medical School Berlin, Faculty for Medicine, 14197 Berlin, Germany

6 Department of Experimental Audiology, Otto von Guericke University Magdeburg, 39120 Magdeburg, Germany

7 Department of Otolaryngology, Head and Neck Surgery, Otto von Guericke University Magdeburg, 39120 Magdeburg, Germany

\* Corresponding Author: [lasse.gueldener@ovgu.de](mailto:lasse.gueldener@ovgu.de); Department of Experimental Psychology, Institute of Psychology, Otto-von-Guericke-University, Magdeburg, Germany,

**Abstract**

Foraging in animals, including humans, confronts the foraging subject with the exploration-exploitation dilemma: continuous exploitation of a given patch leads to the depletion of resources, and the forager must decide when to switch location to find new energy resources. An extensive literature on foraging in humans and several animal species has stated various so-called patch-leaving rules governing this decision. Among others, environmental variations in patch quality (i.e., reward probability) affect the fitness of these rules. In this study, we adopted a probabilistic foraging regime and two comparable tasks designed to study the effects of this regime for patch-leaving behavior in humans and gerbils. Humans were administered a visual search task, whereas gerbils performed a foraging task in which they traveled between two locations at which food rewards were expected. In both tasks, patches varied randomly in their initial reward probability (i.e., patch quality), and the reward probability decayed exponentially mimicking a quickly depleting reward source. Our data showed that patch-leaving decisions of both species followed an incremental mechanism based on reward encounters that is considered optimal for maximizing reward yields in variable foraging environments. The two species differed in their giving-up times but only humans tended to overharvest. However, gerbils and individual humans who did not overharvest were sensitive to declining collection rates in accordance with the marginal value theorem. Altogether this study introduces a novel paradigm for a between-species comparison on how the exploitation-exploration dilemma can be solved. Despite deliberate task-differences the results suggest that the underlying heuristics humans and rodents use to time their patch-departures can align.

**Introduction**

“Should I stay or should I go? “- In natural situations, foraging animals constantly find themselves confronted with the dilemma to either keep exploiting a current source of energy thereby depleting it more and more, or to move on and explore the environment for novel sources of energy. If we picture roaming antelopes, they either keep staying in the same feeding grounds, maybe because there is a water hole nearby. Or, they decide to move on to search for new areas with more fresh grass, yet taking the risk that there may be neither fresh water, nor fresh grass, or more predators around. Similarly, a hunter and gatherer must decide whether to follow the roving antelopes or to stay and wait for new prey to arrive. This scenario shows that foraging choices and strategies affect an individual's biological fitness and that foraging patterns have developed to maximize this fitness (i.e., optimal foraging, Pyke, 1984). It also illustrates that foraging behavior often requires the balancing of conflicting demands (exploitation versus exploration) and that tradeoffs play a crucial role in shaping foraging behavior (Sih, 2011).

Overall, there is no clear rule when to stop exploiting and when to decide to explore. This so-called patch-leaving behavior can be driven by external events as well as by internal urges. Here, we report results from a behavioral study in which we tested 44 human participants in a probabilistic foraging task and compared their patch-leaving behavior to that of 18 gerbils. For this purpose, we designed two distinct foraging tasks suitable for the respective subjects. Yet, the two tasks were similar enough in their principle operationalizations to allow for comparisons between humans’ and rodents’ patch-leaving decisions. Our central goal was to analyze the reward-dependent foraging behavior across the two species using two specific paradigms that can eventually pave the way for more research on cross-species comparisons. Importantly, this does not imply that specific patch-leaving behaviors can be expected to universally occur, independent of the experimental paradigms used.

The human participants performed a visual foraging task embedded in a visual search paradigm (e.g., Kristjánsson et al., 2020; Wolfe, 2013). Animals, including our hunter-gatherer ancestors, often encountered foraging environments with spatially and temporally distributed *patches* (e.g., forest districts with varying prey richness at different locations and distances within one habitat). Foraging in such environments involves *serial* decision-making, incurring temporal travel costs as animals move from one patch to another. It is this very aspect of foraging that is better captured in a serial visual search task compared to the bandit-like gambling tasks, traditionally used to study the exploration-exploitation dilemma (e.g., Daw et al., 2006; Laureiro‐Martínez et al., 2015), where, in the latter, decisions involve simultaneous choices. In our visual search task, the participants searched on monitor displays for target items among distractors and had to decide whether to continue searching in the current display (i.e., patch) or to switch to a new display with novel targets by pressing a button (patch-leaving). In each display, participants earned a monetary reward each time they located a target item using the PC-mouse. This approach resonates with previous work on foraging behavior in humans that also combined a serial visual search paradigm with a patch-based foraging task (Kristjánsson et al., 2020; Wolfe, 2013). What works well in humans can be much more difficult apply in rodents. Thus, to introduce adequate task conditions for patch-based foraging for the gerbils, we adopted an established foraging task from Lottem et al. (2018). Here the foraging setup consisted of a box-like arena with two foraging spouts located on opposite sides of the box. The animals were trained to nose-poke (forage) at one of two spouts that dispensed food rewards). Importantly, both the human and the rodent task were similar in their probabilistic as well as in their patch-based structure: reward probabilities were exponentially decreasing in both paradigms so that the foraging success declined the longer the foraging humans and animals remained in the same patch. Thus, both animal and human subjects were forced similarly to make patch-leaving decisions for optimal foraging.

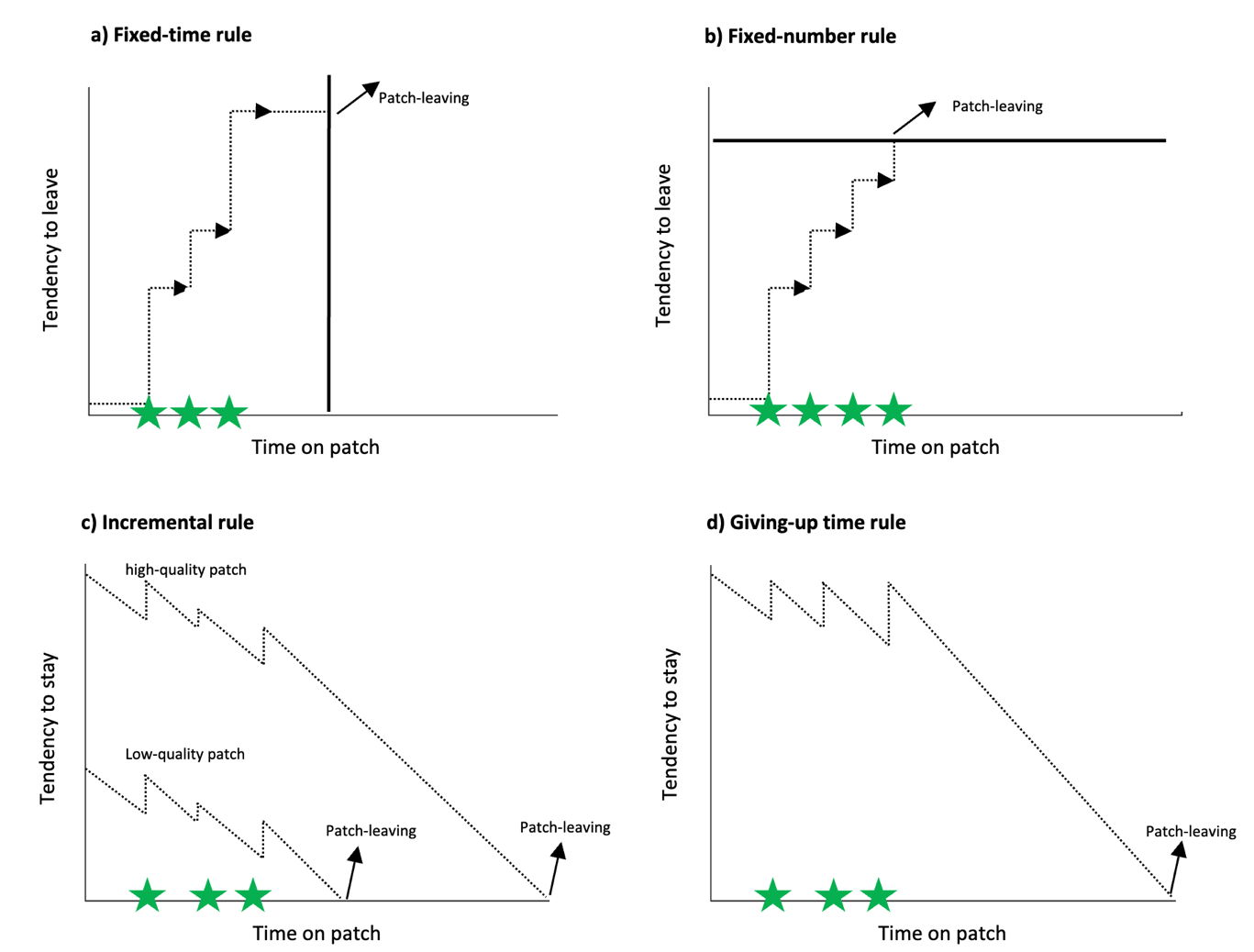
A plethora of theories and formal models have been proposed to model the decision process that results in exploration and to predict the time point of patch-leaving (e.g., Krebs et al., 1974; Charnov, 1976; Green 1984; 1985). All these models share the assumption that an environment’s reward probabilities and variations in patch quality affect patch-leaving behavior and identify decision rules about the optimal time point when to leave a patch (Wajnberg, 2006; Wilke et al. 2009). The key question in our study was to what extent a human and rodent forager can make use of implicit probabilistic knowledge about the environment as well as of their subjective reward histories to make their patch-leaving decisions. One of the most prominent models of optimal foraging (i.e., maximizing the total energy or reward intake during time-constrained foraging) is the *marginal value theorem* (MVT). Introduced by Charnov (1976), this model predicts that a foraging agent should leave the current patch when the instantaneous energy intake rate (i.e., collection rate of energy or reward, ICR) within a given patch diminishes to the average energy-harvesting rate in the entire environment (i.e., the mean collection rate, MCR). This implies that the foraging animal must have access to and use its memory about its recent foraging experience and is able to keep track of its current and average energy intake.

To promote the use of such a probabilistic patch-leaving rule based on reward history, we manipulated the reward probability in two ways. Firstly, in both paradigms, we introduced an exponential decay of reward mimicking a quickly depleting food source (Lottem et al., 2018): once gerbils received a reward after nose poke, the probability of the next reward at the same spout continued to decrease exponentially to zero. Similarly, in the human task, a recently collected target item not only became “inactive” and remained in the search display, but also additional target items were rendered “inactive'', so that a fixation of these deactivated targets would not result in a reward. The number of these incremental target deactivations followed the same exponential decay functions that were used in the gerbil task. This manipulation made the search in each display increasingly difficult over time and was intended to facilitate participants’ tendency to switch to a new display. More importantly, for any given patch, this approach allowed us to avoid reward probabilities following a Poisson distribution. Given a Poisson distribution, the number of preys in a patch is expected to be random and independent of the amount of time that has elapsed since the animal arrived at the patch (Iwasa et al., 1981). In other words, the expected rate of reward on a patch does not decline over time, as it does in environments in which resources deplete as a function of time spent in the patch. Under these conditions, the optimal strategy is to simply spend a fixed amount of time on the patch, regardless of the number of energy sources found, and then move on to the next patch (i.e., ‘fixed-T’ rule, see Figure 1 a)). This strategy maximizes the expected rate of reward per unit time, because the average number of reward items found in a given area or time period is *constant* and *independent* of the time spent searching (see Iwasa et al., 1981). This heuristic does not require knowledge about the subjective average and instantaneous reward intake rate. However, it is no longer optimal once the reward probability depletes as a function of residence time (i.e., the time that is spent foraging per patch) as is the case in our task. In addition, the initial reward probabilities varied randomly between high (100%), medium (75%), and low (50%). Importantly, these patch qualities were not cued initially. This approach prevented single patches from having identical numbers of available rewards that could be received. If the latter was the case, the most optimal patch-leaving rule would be to capture a fixed number of rewards, e.g., 10 rewards, and leave the patch once that number is reached (i.e., the ‘*fixed-N*’ rule, see Figure 1 b)). However, this rule is no longer optimal if patches vary in their quality a lot, and the time to forage is limited (e.g., Wilke et al., 2009). In low-quality patches, foragers using the fixed-N rule would spend too much time to obtain the fixed number of rewards, impairing their overall capture rate given that foraging time is limited. Comparably, using a fixed-T rule, in poor quality patches foragers would have fixed time intervals with few or even zero reward captures as they would continue foraging stoically until the fixed residence time is reached. Thus, environments with variable patch qualities require a different behavioral adaptation for optimal foraging.

In highly variable environments it is rather difficult to reliably estimate the quality of the current patch at the time of entering, but each reward capture is an indication that the current patch is of high quality, thus, increasing the tendency to stay (see Wilke, Hutchinson, Todd & Czienskowski, 2009). This would result in an *incremental rule* (Figure 1 c)) according to which the probability to remain in a given patch starts to continuously decrease with arriving at a new patch but each reward capture increases the probability to stay (e.g., Waage, 1979). Using this rule, foragers do not have to initially estimate the patch quality which would be difficult or impossible, but can mostly rely on foraging success to make this estimate. Foragers using this rule should spend more time in high- compared to medium- and low-quality patches. Moreover, should each novel reward capture increment the forager’s residence time independent of the quality of the current patch, that is, regressing residence times on the number of reward captures on subject-level should result in positive slopes.

Similar to the incremental rule, also the *giving-up time rule* (GUT rule, Krebs et al., 1974; McNair, 1982, see Figure 1 d)) does not require a prior judgment/knowledge about the patch-quality. The GUT rule states that a forager only tolerates a certain amount of time without a new reward capture since the last reward capture. Once this temporal threshold is exceeded, the forager leaves the patch. Each new capture, on the other hand, resets the tendency to stay. One could think of a countdown timer that starts to count down as soon as the forager enters a patch and now runs down steadily. Every reward capture resets and restarts this countdown timer. If no reward is captured before the timer expires, the forager leaves the patch. Thus, in rich patches (i.e., high reward probability), prey is encountered more frequently, and the countdown timer is reset each time so that the forager using a GUT-rule is predicted to spend on average more time in high- compared to lower-quality patches. At the same time, GUTs should be constant within individuals and should exceed the durations of individual’s intervals between two target captures (inter-target intervals; ITI): if a subject has a GUT threshold of 4 s, then her ITI should always be shorter or equal to 4 s, because she leaves the patch in the moment the ITI is about to exceed the 4 s threshold. Similar to the incremental rule, the GUT rule makes use of the past success rate to estimate the upcoming success rate and does not require a prior judgment of the patch quality. As a result, the GUT rule still guarantees fitness if the environment contains patches that vary widely in quality and if the patch quality is difficult to assess in advance (Wilke et al., 2009).

Given the quick depletion of reward and the unpredictable variations of patch-quality, we hypothesized that humans and gerbils are sensitive variable reward probabilities and adopt either an incremental or a GUT-rule to optimize their patch-leaving decisions. This assumption/hypothesis is consistent with previous findings in mice tested in the same foraging task as our gerbils (Lottem et al., 2018). Although the study by Lottem and colleagues did not report GUT data, the mice’ patch-leaving behavior was best explained by a model in which reward captures incrementally increased the probability to stay, analogous to the incremental rule. The authors also reported evidence in support of the MVT in that the mice’ ICRs at the time of leaving were almost identical and statistically indistinguishable to the MCR. In line with this finding, we also hypothesized that ICRs the gerbils approximate the MCRs at the time of leaving. Previously, humans performing a visual search based foraging task did not conform to this prediction of the MVT. They tended to stay in a patch longer than expected. Yet, in contrast to the present study, the foragers were given the choice to switch between target types within a given patch, whereas in our task subjects searched for the same target item throughout the entire experiment. Particularly when switching between target types, the ICR fall below the MCR, but the subjects stayed in the same display - inconsistent with the central prediction of the MVT (Kristjánsson et al., 2020). Other foraging experiments based on visual search paradigms showed deviations from the MVT's projections when the patch quality varied largely (i.e., when subjects foraged in patches that had one out of ten randomly chosen reward probabilities) as well as when visual information was reduced to the extent that foragers could no longer discern whether a target item was associated with a reward (Wolfe, 2013, experiments 5 and 6). This evidence suggests that rendering foraging tasks more complex, patch-leaving behavior appears to be no longer in accordance with the MVT. Compared to these studies, the human task reported here was less complex: no switches between targets needed to be regarded as part of the foraging strategy and the underlying reward probability structure (i.e., patch quality) of the entire environment, although changing randomly from patch to patch, was less variable. Given this simpler task conditions that matched those of the gerbil task, we expected that also humans’ ICR to approximate the MCRs at the time of leaving in keeping with the MVT (see also Wolfe, 2013, experiment 1).



**Fig. 1. Simple heuristics to time patch-leaving decisions**. Using a fixed-time rule, the patch is left independent of the number of prey encounters (green stars) (a), whereas a patch is left after a fixed number of prey encounters have been found if a fixed-number rule is used (b). According to the incremental rule, each prey capture increases the probability to stay in a patch postponing the patch-leaving (c). Using the giving-up-time rule, the tendency to stay in the patch declines as a function of unsuccessful search and each prey capture resets it to a maximum. Adopted from Wilke et al., (2009).

**Methods**

**Human Study**

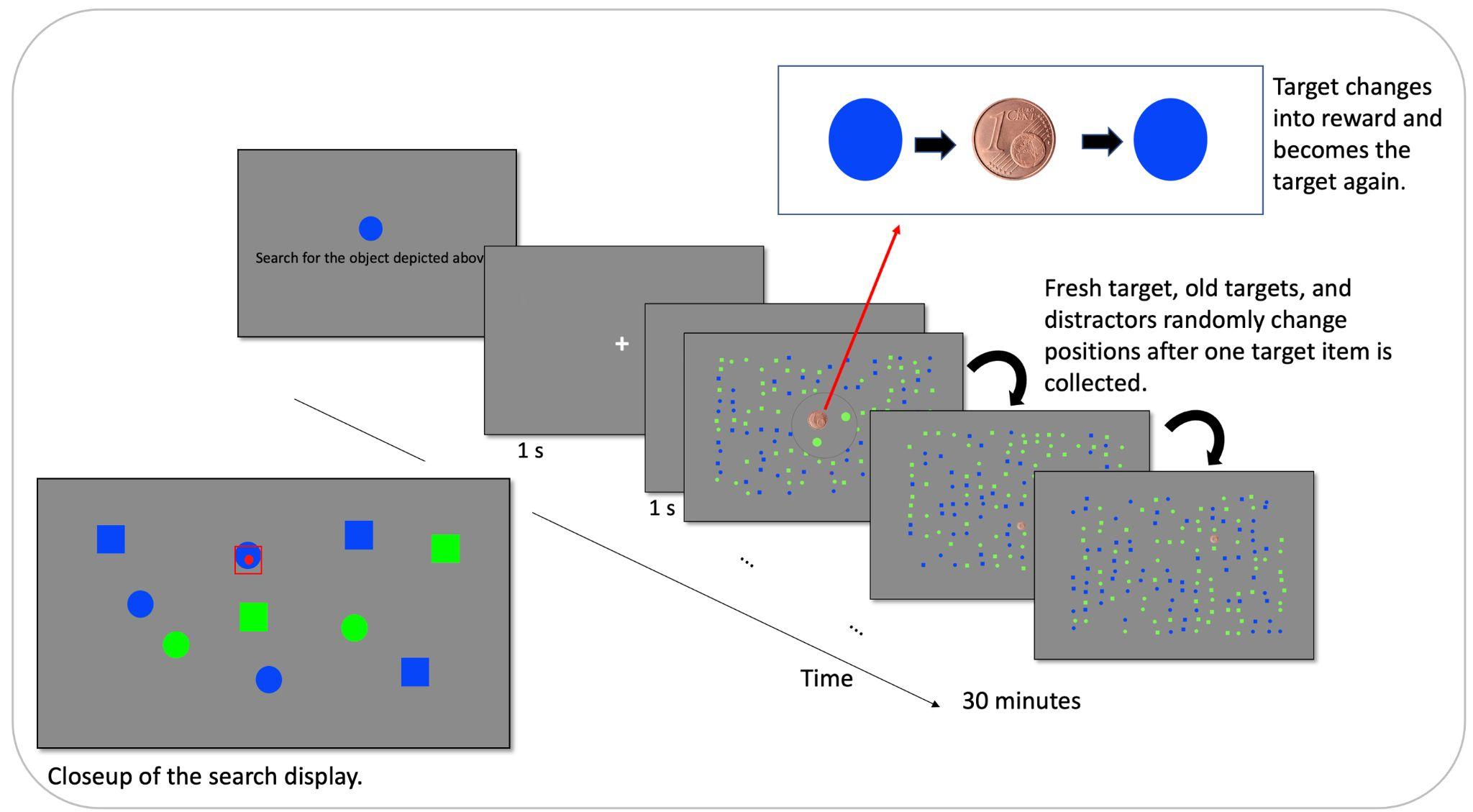
**Participants.** 52 (17 male) native German speakers participated in the experiment. 32 of the participants were tested in the PC-laboratory (lab), while 20 subjects were tested in a follow-up neuroimaging experiment using functional magnetic resonance imaging (fMRI-lab). For this report, we used the behavioral data of the fMRI subjects to increase the original sample size. All volunteers were between 19 and 37 years old (*M* = 24.25 years), right-handed by self-report except for two participants, and had a normal or corrected-to-normal vision. They provided written consent and were monetarily reimbursed based on the earnings they made performing the foraging task. We excluded eight PC-lab participants from the data analysis because they performed less than six trials in at least one of the reward conditions. Another two PC-lab subjects who yielded an overall number of trials below the 1st or above the 99th percentile of the group distribution were also excluded. Thus, the final sample size was 44.

**Visual search paradigm**

**Set-up and Stimuli.** We used the Python toolbox “PsychoPy” (Peirce, 2019) to control the stimulus display and responses. The stimuli were presented on a 24′′ Samsung monitor (1920:1080 resolution, 60 Hz refresh rate). All participants were positioned 50 cm away from the screen. Stimuli consisted of geometrical forms. These were either squares or circles that appeared in either blue or green color. All stimuli subtended 0.59° visual angle. Their spatial locations were randomly assigned on a spatial grid spanning a rectangle field of 12.9° \* 14° visual angle. One stimulus type (e.g., all blue circles) was assigned as targets while the three remaining stimulus classes would serve as distractors. As reward indication, we used an image of a Euro-Cent symbol that subtended 0.70° visual angle.

**The experimental task.** To study the exploration-exploitation dilemma in human subjects we designed a probabilistic foraging task. Participants were asked to search and collect target items among distractors in a visual feature conjunction search task. Stimuli consisted of simple square- and circle-shaped objects randomly located in the search display. They used the mouse to navigate through the display. Target items were defined by a specific conjunction of shape (i.e., circle) and color (i.e., blue), and equal numbers of distractors would differ either in shape (blue squares), in color (i.e., green circles), or in both feature dimensions (i.e., green squares). The total search time was restricted, and a countdown timer was constantly visible to the participants at the left bottom corner of the display. To obtain a reward, participants had to navigate the mouse pointer to a target. Once a target had been fixated for 300 ms, the target turned into a reward indicator (i.e.,1 Euro-Cent) and then returned to its previous appearance. This served as the feedback that the target had been "foraged”, and a reward was received. The participants were then able to continue the search for the next target in the display. At the display’s left bottom corner, the participants were able to constantly track the total number of rewards they had already earned. At the display’s right bottom corner, they could keep track of the remaining time.

Once a collected target returned to its previous appearance, it was turned inactive so that a second fixation of the same target would not result in a further reward capture. Moreover, with each collected target item, additional targets, randomly located in the display, were also deactivated. In this way, after a fixation on a new target, the remaining reward probability in a given display decreased exponentially, mimicking a quickly depleting food source. Furthermore, the whole spatial configuration of target and distractor locations changed after each collected target, and all targets and distractors consequently appeared at new locations. This manipulation made the search increasingly difficult and quickly inefficient as it was impossible to remember target locations that had been already visited and, thus, would not promise a new reward following fixation (Horowitz & Thornton, 2008). To compensate for this, participants could choose to end the search in a display and to proceed to the next display at any time. The countdown did not pause when volunteers were directed to the next display, and each switch to a new display consumed time (3.5 s in the PC-lab, ~ 5.75 s in the fMRI-lab; see Task design) analogous to patch-leaving costs in ecological foraging or the movement from one foraging spout to the other in the gerbil experiment. After volunteers pressed the spacebar to continue to a new display, a central fixation cross appeared for one second followed by a blank screen for another second before the new display appeared. With the appearance of a new display, the fixation cross (mouse point) was relocated to the display center. A depiction of a trial sequence is shown in Figure 2.



**Fig. 2: Human visual search task.** Diagonal sequence represents a trial. At the beginning of a session, the target object (here a blue circle) was introduced. The beginning of a trial was cued by a central fixation (1 s) followed by a blank (1 s). Next, the search display (i.e., patch) appeared. By navigating the mouse cursor to a target (red square with red dot at center in lower left), participants realized a reward capture. Upon such a capture, the collected target turned into a reward for 500 ms (1 euro-cent image) and then changed back to its previous appearance. An already collected target would not turn into a reward again if fixated again. With each reward capture, all items changed positions randomly. At any time, participants were able to switch to a new display by button-press.

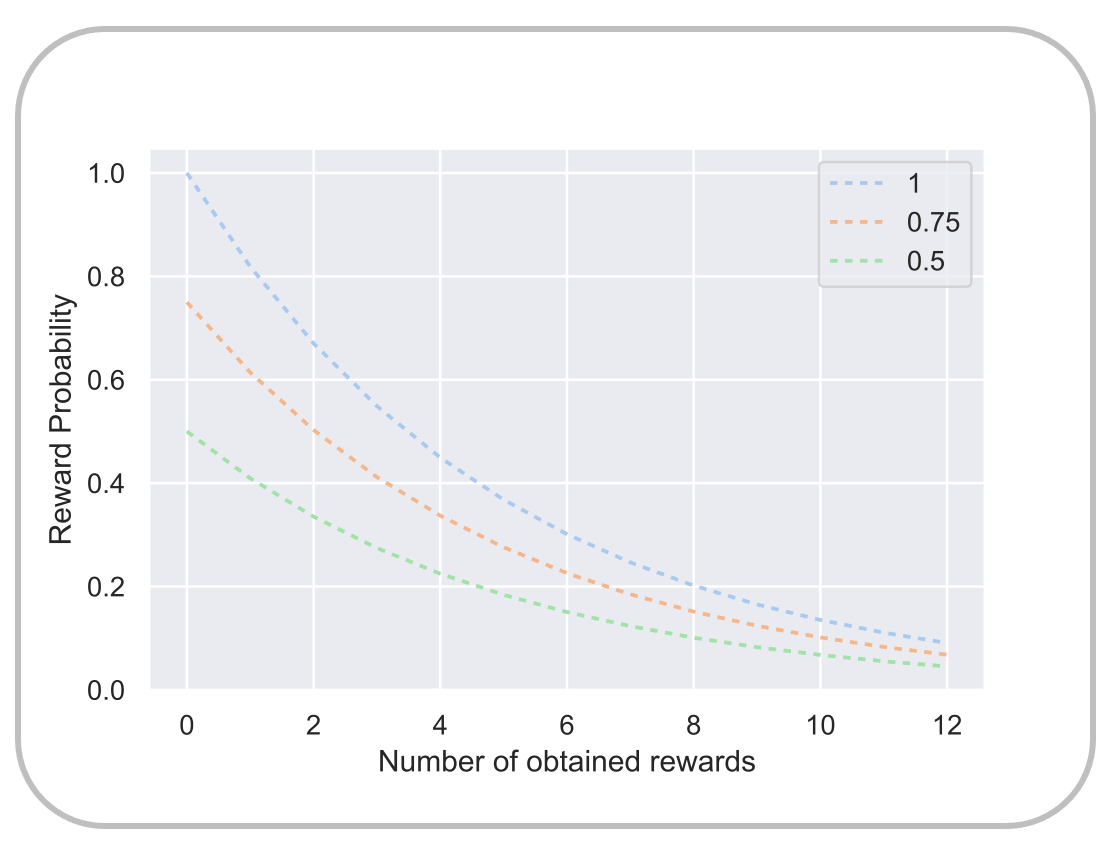
**Task Design.** We aimed at determining the optimal task conditions to study patch-leaving behavior in humans. This required a task in which volunteers actively decide to leave an exploited patch and to switch to a new patch. This exploratory foraging behavior is facilitated if the overall number of search targets in each search display is high and if the travel costs moving from one display to another is relatively short (Cain, Vul, Clark & Mitroff, 2012; Wolfe, 2013). Therefore, we chose a total number of 40 targets and a relatively short travel time of 3.5 s. Due to additional intermediate data storage during the travel in addition to generating the upcoming display, travel times in the fMRI experiment were on average 5.75 ± 0.1 s.

To make reward encounters dependent on the foregone foraging success and, thus, mimicking a depleting patch, we chose an exponential decay function for the depletion of reward following each reward capture (Lottem et al., 2018). That is, the number of available rewards drastically decreased within a relatively short period of time, resulting in an inefficient search. Due to the time constraint, participants decided to switch to a new display to improve search efficacy. The following decay function was adopted from Lottem et al. (2018) (see Fig. 3):

1. P(on = 1 | ti) = Aie (-(n-1))/5

here *ti* is the *i*th trial type, i.e., low-, medium-, and high-quality trials. These trial types had different exponential scaling factors A1 = 0.5, A2 = 0.75, A3 = 1. *N* indicates the number of already achieved reward captures (previous target fixations that resulted in an earning) within a trial. *On* is the positive outcome of the *n*th target fixation (1 for reward). We additionally varied the initial reward probabilities by applying three conditions from high, middle to low probabilities to be able to test whether subjects adapted their patch-leaving strategies according to probabilistic changes in the environment. In the high reward condition (i.e., high quality patch, blue function in Figure 3), all 40 target items (100%) were active and would turn into a reward following a first fixation. In the medium reward condition, only 75% of all targets (i.e., medium quality patch, orange function in Figure 3) were active and were associated with reward following a fixation. Only 50% of all targets were initially active in the low reward condition (i.e., low quality patch, green function in Figure 3).

The participants started the experimental session with two training trials in which they searched for target items in two consecutive displays without the time constraint, and no reward was registered. Once volunteers terminated the search in the second display, they were informed that the main experiment would start next and that they were given a total search time of 30 minutes (PC-lab), or 6 x 10 minutes (fMRI-lab), respectively. The fMRI session took place on a single day and participants were able to take short breaks between the runs remaining in the scanner.



**Fig. 3**: **Reward probability protocol**. Graph shows the decaying reward probability as a function of obtained rewards in a given trial (i.e., patch) for the three reward conditions (100%, 75%, 50% initial reward probabilities – include label into legend). The exponential decay function was adopted from Lottem et al., 2018.

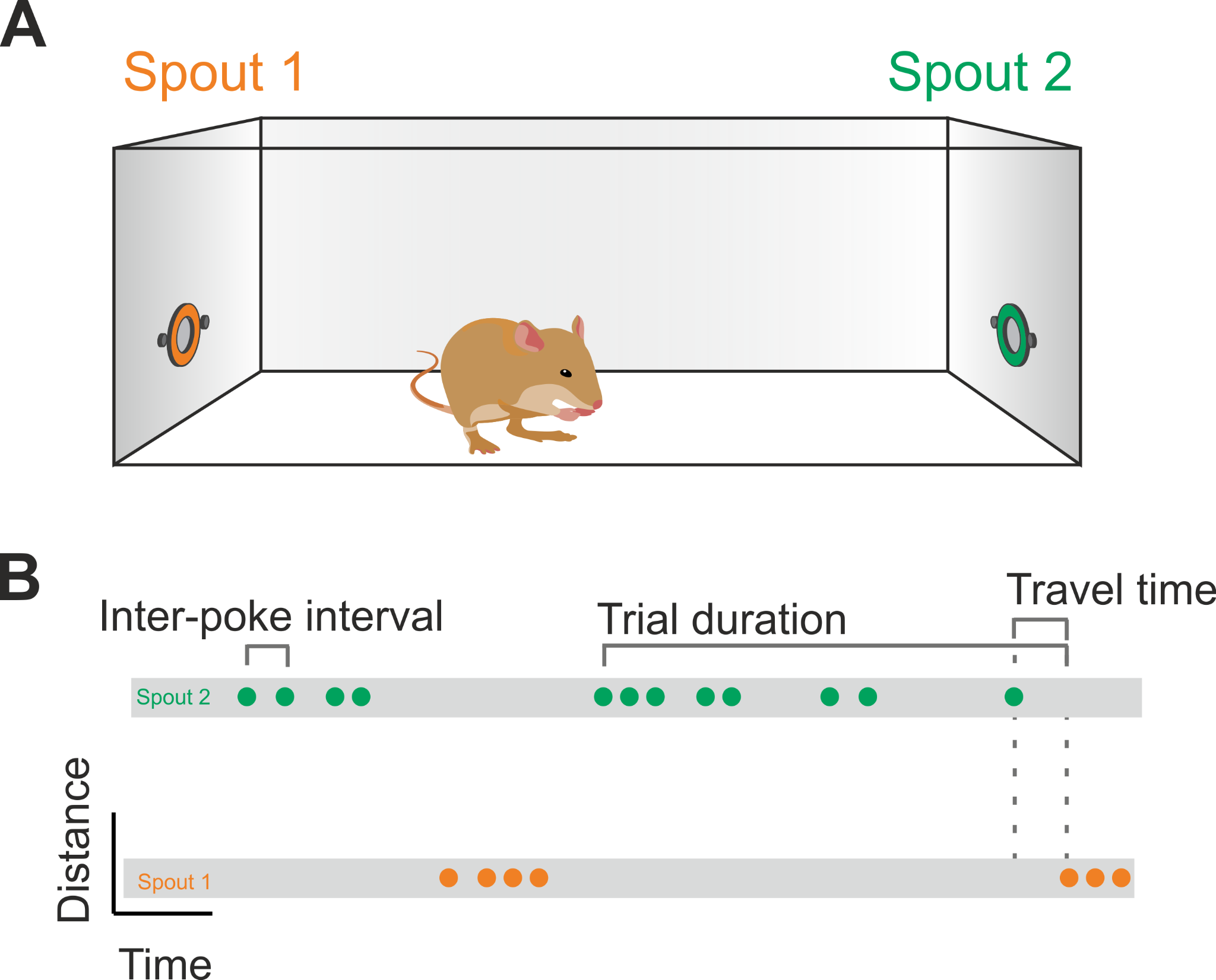
**Animal study**

Animal experiments were performed with 18 adult male Mongolian gerbils (in-house bred). The age of the animals during these experiments varied between three to four months. All experiments were performed in accordance with the German animal welfare law (NTP-ID: 00041189-1-X.).

**Food restriction.** The animals had free access to water but were food restricted starting three days before the beginning of the foraging task. Before food restriction was started, the animals’ body weight was measured over three days to obtain an averaged baseline body weight (BBW). The BBW of the animals was 70-80 g before starting the foraging task. To keep the animals' body weights above the critical level (85% BBW) during the foraging task period, food was supplemented inside the cage at least 2 hours after the end of the foraging task. The total daily food intake including the amount of food retained during the trainings session of an animals was between 3-7 g based on the performance of the animals in the foraging task**.**

**Foraging setup and stimuli.** Foraging tasks were performed in a ‘foraging box’ placed inside an electrically shielded and sound-proof chamber. The box had a wooden framework [MOU7] , and the walls consisted of vertical cylindrical plastic bars placed 1 cm apart from each other. The floor of the box consisted of a plastic mesh. Two foraging spouts were placed on opposite sides of the box and attached to food dispensers (Campden Instruments Ltd., USA). The dispensers were operated by custom-built Arduino hardware that was controlled by a custom-written application program in MatLab (Version 2019). On the sides of each spout, an infrared sensor pair was located to register the nose-pokes of the animals. The dimensions of the foraging box were 37 cm x 26 cm x 48 cm. The distance between the spouts was 36 cm. The foraging setup for the animals is shown in Figure 4.

**Probabilistic foraging paradigm.** Each animal was trained once per day. In each training session, the number of trials was dependent on the animals’ behavior. The total foraging time was restricted to a maximum of 30 minutes during the initial learning phase of five sessions. After 3 training sessions, the animals mastered the task more quickly, and a single experimental session was typically concluded after 15-20 minutes once the animal became disengaged. Each animal performed 20 sessions on 20 consecutive days. A trial was defined as the time between entering a given spout and switching to another spout. In every trial, we recorded the number and durations of nose-pokes at the spouts. Nose-pokes with a duration of less than 100 ms were counted as errors, pokes lasting longer than 100 ms were recorded as hits. Error pokes remained unrewarded. Hit pokes were either rewarded with 20 mg of commercially available food pellets (Dustless precision pellets, Grain based, 20mg, Plexx B.V.) or unrewarded based on the current reward probability and reward outcome. The reward probabilities decreased with increasing numbers of pokes following the same exponential decay function deployed in the human visual search task (see Figure 4). As a result, the probability of new reward capture diminished quickly, encouraging the animal to alternate between the spouts during the foraging session. Like in the human task, we used three different patch qualities (100%, 75%, and 50% reward probabilities) that were randomly interleaved between consecutive trials. To obtain more trials from the animals and to maintain the motivation of the animals, the reward probability was set to zero after the 20th hit nose-poke of a trial. After each reward, a dead time of 100 ms occurred.



**Fig. 4.: Probabilistic foraging task in Mongolian gerbils.**  **A)** Schematic illustration of the foraging task in which gerbils can access food from two ‘Foraging spouts’ 1 and 2 located at the opposite ends of the box. **B)** Scheme indicates the animals’ foraging behavior as a function of time. The spouts appear at the illustration’s top and bottom; the gray bars both indicate the corresponding pokes at each bar over the timeline of the experimental session. Traveling between the two spouts, the green and orange dots indicate nose pokes performed at the respective spout. Trial duration: A trial started with the first breaking of the light barrier by a nose poke lasting more than 100 ms (hit) at one spout and ended with a first poke at the opposite spout, as indicated in the figure so that the residence time per spout was given by the time interval between trial start and end. This last release of a poke in each trial marked the beginning of the following travel time that corresponds to the time until the first hit poke at the opposite spout started.

**Statistics and Reproducibility**

We used custom-written code in Python (version 3.6) to perform all data analyses. First, we aimed to test whether the foraging behavior in gerbils and humans was modulated by differences in the patch quality and whether these modulations differ between the two species. Potential modulations of the foraging behavior by patch quality would indicate that humans and gerbils adjust their patch-leaving strategies based on probabilistic differences in the reward in a given environment. To test this, we used single factor repeated-measures analysis of variance (ANOVA), as well as mixed ANOVAs with Tukey HSD correction for multiple comparisons using the Python package “pingouin” (Vallat, 2018). QQ-Plots were used to screen for violations of normality in which case we used non-parametric measures (e.g., Friedman tests with Nemeny post-hoc contrasts). Greenhouse-Geiser corrections were applied in case of non-sphericity. In all analyses, the averaged medians were used without outlier correction to report descriptives as well as for statistical testing. For within-subject regressions we used the ‘lingress’ function from python’s SciPy (version1.10.1) statistics. Cox-regressions were performed using python’s lifelines package.

**Results**

**Number of reward captures and residence times increased with patch-quality**

Timing patch-leaving decisions by a fixed number of reward captures (i.e., the fixed-n rule) would result in equal numbers of rewards- across patches. Inconsistent with this, our human foragers showed an increased number of reward captures with increasing patch quality, *F*(2,82) = ﻿449.848. The averaged number of reward captures did not differ between humans tested in the PC laboratory and those tested in fMRI lab, *p* = ﻿0.259. The highest number of rewards was yielded in high-quality patches [*M* = 11 ± 3], a lower number of rewards was captured in medium-quality patches [*M* = 8 ± 3, *t*(40) = ﻿-4.606, *p* = .001], and the lowest number in low-quality patches [M=5 ± 2, *t*(40) = ﻿-10.196, *p* = .001].

The same pattern of results was observed in the foraging gerbils. They also achieved more reward captures with increasing patch quality: *F*(2,34) = ﻿1052.868. The reward yield in low-quality patches was 2 ± 0.5 rewards on average, in low-quality patches, 3 ± 0.5 in medium-, and 4 ± 0.5 in high-quality patches. The differences were statistically significant between high- and low- [*t*(16) = ﻿- ﻿22.095, *p* < .001], and high- and medium-quality patches [*t*(16) = ﻿-10.240, *p* < .001]. Point plots of averaged median rewards obtained as a function of patch quality (i.e., start reward probability) for both species are shown in Figure 5 a).

The ‘fixed-time’ rule states that a forager would spend an equal amount of time in a patch regardless of the current intake rate or given patch quality. Thus, one would expect equal residence times across all three patch qualities. Trial durations did not differ between both human samples and were again averaged for the analysis, [fMRI: *M* = ﻿41.992 ± ﻿17.330; PC: *M* = ﻿37.714 ± ﻿17.617, *t*(40) = ﻿0.773, *p* = .221]. Unlike the prediction of the fixed-time rule, humans’ averaged median residence times were modulated by the reward probability, *F*(2,82) = ﻿66.955, *p* < .001, with longest average residence times of ﻿45.594 ± 17.506 s in the high quality patches, followed by 40.561 ± 17.453 s in medium-quality, and the shortest residence time [33.100 ± 19.100] in low-quality patches. Yet, post-hoc tests showed that only residence times in high quality patches were significantly higher than those in low quality patches, *t*(41) = ﻿-3.132, *p* = .006; low vs. medium: Tukey's HSD *p* = .152, medium vs. high: *p* = .421 ).

Similarly, gerbils' residence times increased with patch quality, *F*(2,34) = 171.259, *p* < .001. The averaged median time that gerbils spent in the low quality patches [﻿9.789 ± ﻿1.884 sec] was on average 3.30 ± 0.795 (SE) s shorter than the averaged median time spent in medium quality patches [﻿13.627 ± ﻿2.110 s, *t*(16) = ﻿-4.356, Tukey's HSD *p* = .001], and ﻿-7.879 ± 0.881 s less compared to high-quality patches [17.668 ± ﻿3.435 s, *t*(16) = ﻿-8.942, *p* = .001]. Averaged residence times as a function of patch quality (i.e., start reward probability) for both species are shown in Figure 5 b).

Given the rapid depletion of rewards, subjects were encouraged to readily leave a current patch instead of spending too much time in it. Consistent with this, human residence times were negatively correlated with the total number of reward earnings (in €) they yielded in an entire foraging session [PC-lab: *r* Pearson= -0.717, *p* < .001; fMRI-lab: *r* Pearson= -0.653, *p* = .002] (see Figure 5 c) and d)). In other words, the more time subjects invested searching per patch, the less earnings they yielded throughout the entire experiment. No evidence for such a relationship was found in the gerbil data, [*r* Pearson= 0.280, *p* = . ﻿258].

**Ein Bild, das Text, Reihe, Diagramm, Screenshot enthält.

Automatisch generierte Beschreibung**

**Fig. 5**: **Task performance measures. a)** Dot plots show the number of reward captures as a function of patch quality for humans (dark red) and gerbils (orange). Small dots indicate individual subjects’ values, green diamonds the sample mean. **b)** Residence times for humans (dark red) and gerbils (orange) as a function of patch quality. **c)** - **e)** Plots show the relationship between the total number of rewards obtained at the end of foraging sessions and residence times per patch. In both the PC-lab (**c**) and the fMRI-lab (**d**), humans’ residence times were negatively correlated with total earnings. The association was reversed in gerbils (**e)** but was not statistically significant.

**Splitting groups by their median giving-up times**

We observed a large variation in human participants’ residence times [range = ﻿66.936 s] and giving-up times [GUTs; range = 14.419]. This was in stark contrast to the very consistent gerbil data (range of residence time = ﻿7.866 s; range of GUT = ﻿3.580). In human participants, residence times strongly correlated with the GUTs, *r* Pearson = 0.729, *p* < .001. Like human residence times, GUTs of both the PC- and the fMRI-lab samples correlated negatively with total earnings, *r* Pearson = -0.668, *p* < .001, *r* Pearson = -0.676, *p* = .001. No such correlation was found in gerbils, *r* Pearson = ﻿0.350, *p* = . 153

Thus, to better account for the heterogeneity in humans, we split the group by its median GUT (﻿6.951) into a long- and a short-GUT group [long-GUT, n = 20: residence time: *M* = 49.960 ± 14.729 s, GUT *M* = 10.414 ± 2.036 s; short-GUT, n = 20: residence time *M* = 30.415 ± 14.729 s, GUT *M* = 4.138 ± 1.842 s]. Both residence times and GUTs for both subgroups are shown in Figure 6 a) and b). Unsurprisingly, the long-GUT subjects had significantly longer residence times compared to short-GUT subjects, *t*(42) = 4.412, *p* < .001. GUTs were on average significantly longer in the fMRI subjects, [fMRI: M = ﻿8.510 ± ﻿3.543; PC: M = ﻿5.858 ± ﻿3.378; *t*(40) = ﻿2.414, *p* = .021. This means that more fMRI subjects entered the long-GUT group (15 out of 21 subjects), while more PC-lab participants were included in the short-GUT group (14 out of 21). Yet, importantly, *within* the short- and long-GUT groups there were no differences in GUTs between fMRI and PC-lab participants, [long-GUT, *p* = .145, short-GUT, *p* = .150].

For comparison, we also split the group of gerbils in the same way into a long-GUT, [GUT *M* = ﻿5.690± ﻿0.780 s, residence time *M* = 15.488 ± 1.781 s], and a short-GUT group [GUT *M* = ﻿4.347± ﻿0.390, residence time *M* = 11.901 ± 1.012 s]. Also in gerbils, residence times were significantly longer in the long-GUT group, *t*(16) = ﻿4.946, *p* < .001, but in contrast to humans, the two groups of gerbils did not differ in the total reward captures, *p* = .722. Figure 6 c) and d) show the gerbils’ residence times as well as GUTs as a function of patch quality for both groups.

Ein Bild, das Text, Screenshot, Diagramm, Zahl enthält.

Automatisch generierte Beschreibung

**Fig. 6: Residence times and giving-up times as a function of patch-quality after group-splitting.** **a)** Point plots show humans’ residence times as a function of patch-quality. Small dots in orange indicate individual data points of the long-GUT humans, dark-red dots index individual data points of the short-GUT group. Diamonds index the mean values. **b)** GUTs of humans as a function of patch-quality. **c)** Gerbils’ individual residence times after group splitting. **d)** GUTs as a function of patch-quality are shown for the long- (orange) and short-GUT (dark-red) gerbils.

**Prolonged giving-up times indicate a bias for exploitation in humans**

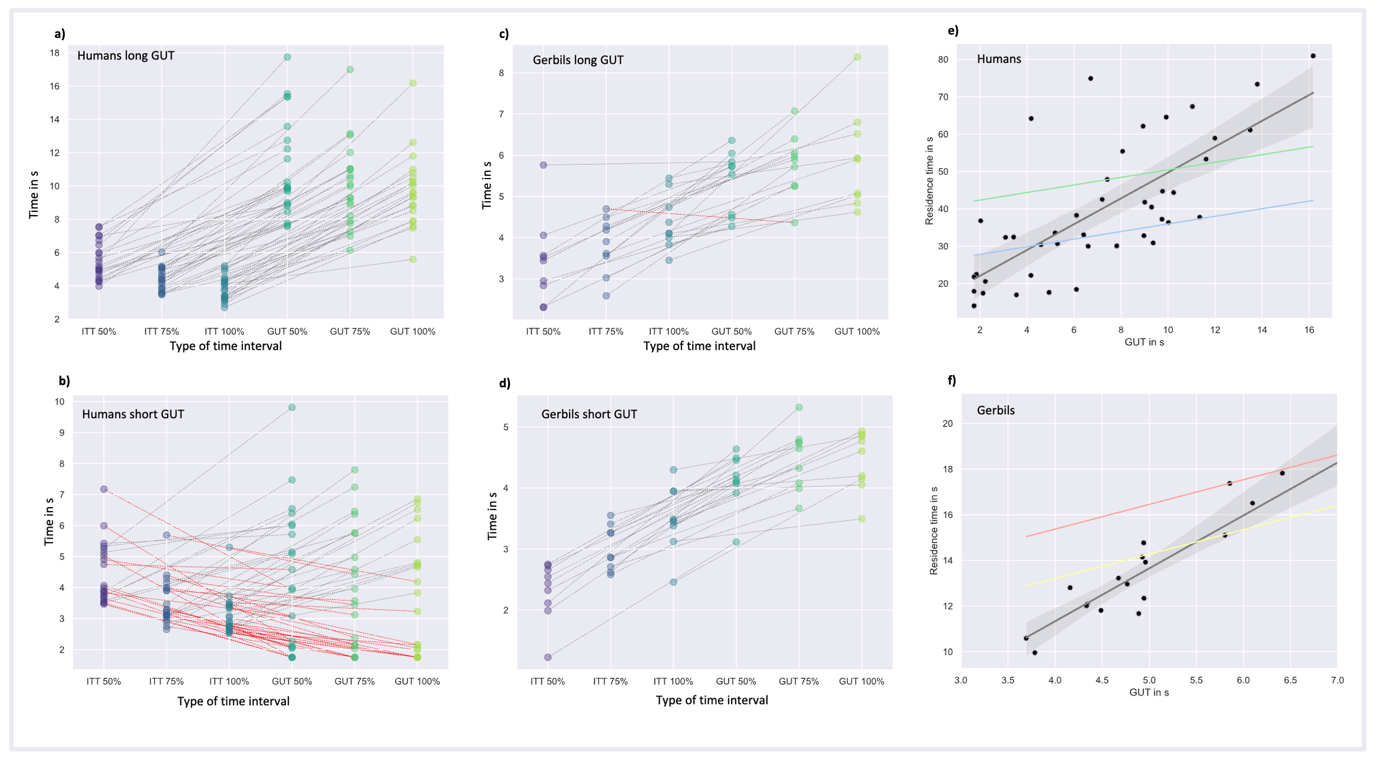
An optimal GUT rule should account for differences in patch-quality with longer GUTs in better patches (McNair, 1982). In contrast to this prediction, already Figure 6 b) shows that the long-GUT humans invested the longest GUTs with an average of ﻿11.033± ﻿2.896 in low-quality patches. Compared to this, GUTs decreased to ﻿10.006 ± ﻿2.422 in medium-, and to ﻿9.699 ± ﻿2.127 s in high-quality patches. Yet, a one-way ANOVA provided no evidence for true differences, Friedman *F*(﻿1.905, ﻿38.095) = ﻿2.674, *p* = .084. In the short-GUT group of humans, average GUTs were ﻿4.250 ± ﻿2.246 s in low-, ﻿3.959 ± ﻿1.940 s in medium-, and ﻿3.769 ± ﻿1.840 s in high-quality patches. This time, we found a significant effect of patch-quality on GUTs, Friedman *F*(﻿1.905, ﻿38.095) = ﻿﻿3.333, *p* = .049. Post-hoc tests provided anecdotal evidence for different GUTs between high- and low-quality patches, *p* ﻿= .054

GUTs with comparable durations across patch types would still be in accordance with a simple GUT rule that does not account for differences in patch-quality (Krebs et al., 1976). If participants applied such a rule, their GUTs should then consistently exceed previous inter-capture intervals (ICI). This is because participants using a fixed GUT to time the patch-leaving, would leave a patch before their ICIs exceeds their GUTs because they only tolerate the fixed duration (i.e., the GUT threshold) without a new capture. Averaged ICIs and GUTs as a function of patch-quality are shown in Figure 7 a) for each long-GUT subject and in b) for short-GUT subjects. In the long-GUT group, all 21 subjects had averaged GUTs consistently longer than their average ICIs across all three patch types. This proportion dropped significantly in the group of short-GUT subjects with only 6 participants displaying a GUT-ICI pattern in support of a simple GUT rule, proportions z-test: *z* = ﻿-4.582, *p* < .001. This pattern of results showed that long-GUT subjects’ behavior was consistent with a fixed GUT rule in all participants of that group. Yet, their GUTs were prolonged to an extent where it affected the overall task performance negatively. In the humans’ short-GUT group only less than half of the participants’ data was in support of a GUT rule. Inspecting single gerbils’ GUT-ICI patterns showed that in the long-GUT animals eight out of nine individuals had on average longer GUTs than ICIs, and in the short-GUT animals, it was only one individual that had average ICIs exceeding its GUTs in medium quality patches (see Figure 7 c and d). Thus, in both groups of gerbils the data pattern was consistent with a simple GUT rule with no significant changes in proportions of animals behaving against the rule’s prediction, *p* = .303.

**Increments in GUTs by residence times in both species**

It appeared that humans tended to prolong their GUTs with increasing trial durations contradicting optimal foraging behavior (Figure 7 a)). Consistent with this notion, within-subject regressions confirmed that humans incremented their GUTs with each additional second spent in the current patch. In both human GUT groups, we observed positive slopes for 19 out of 21 participants, [short-GUT: mean slope of = ﻿1.020 ± ﻿0.715, *t*(20) = ﻿6.374, *p* < .001; long-GUT: , mean slope = ﻿1.013 ± ﻿1.020, t(20) = ﻿4.126, *p* < .001], indicating that participants in both groups extended their GUTs incrementally with each additional second they spent in the current patch. There was no evidence for difference in the slope’s steepness between the two human GUT groups, *p* = .﻿508. The same pattern of results we observed also in the long- [mean slope = ﻿1.082 ± 0.084, *t*(8) = ﻿36.443, *p* < .001] and short-GUT groups of gerbils [﻿1.070± ﻿0.091, *t*(8) = ﻿33.416, *p* < .001] (Figure 7 b)).

Taken together, on average all but one gerbil’s GUT data was consistent with the predictions of a simple GUT rule. In humans, only the long-GUT group behaved in accordance with the simple GUT rule. Yet, unlike the animals, these human subjects seemed to choose suboptimal GUT durations. Given the task conditions, especially due to the quick depletion of reward, there was no benefit in prolonging residence times after the first few target encounters. Long-GUT humans who did this regardless showed significantly poorer task performance compared to short-GUT humans. This difference in performance did not exist between long- and short-GUT animals likely due to the only marginal difference in GUTs between the two groups of animals. Yet, both species showed a tendency to factor in sunk cost into their patch-leaving decision as they extended their GUTs in the current display depending on how much time they had already spent in that patch.

****

**Fig. 7: Testing the GUT-rule. a)** Point plots show individual humans’ GUTs and averaged ICIs in seconds (s) plotted for each patch quality of the long-GUT group. Connecting lines indicate the values that belong to the same individual. In all subjects GUT durations consistently exceeded the average ICIs, in accordance with a GUT rule used for patch-leaving. **b)** Corresponding GUT and ICI data shown for the short-GUT group, red lines show a downward trend indicating 14 subjects who had lower GUTs than ICIs and, thus, a GUT-ICI pattern not conforming to the GUT-rule. **c)** The point plot shows the same GUT-ICI relation for the long-GUT gerbil, the single red line indicates the deviation from the GUT-rule in a single gerbil. **d)** Point plots show short-GUT gerbils’ GUT-ICI patterns. Here, all animals showed data were consistent with the GUT rule. **e)** Linear relationship between residence times and GUT durations for the human sample on group level. The green line represents the average regression line derived from within-subject regressions, predicting GUTs based on residence times within the long-GUT group. The blue line shows the same for the short-GUT group of human subjects. **f)** Linear relationship between residence times and GUT durations in gerbils. The red line equals the average regression line of all gerbils in the long-GUT group obtained from the within-subject regressions (GUT regressed on residence time), and the yellow line shows the same for the short-GUT group.

**Increments in residence times following reward captures in both species**

So far, the data had provided good evidence for the GUT rule in both groups of gerbils as well as in the long-GUT human subjects. The short-GUT group of humans, however, showed data that were inconsistent with such a rule. Thus, we next examined the relationship between reward captures *within* a patch and residence times. Given the unpredictable changes of patch-quality, rewards encountered within a patch provide the only viable estimate of the underlying patch-quality. If foragers relied on this estimate, they would extend their residence times incrementally with each novel reward capture because each new reward encounter would suggest that the current patch is potentially of high quality. To test this, we again calculated within-subject regressions but this time we regressed the residence times on the number of reward captures (Hutchinson, Wilke & Todd, 2008; Mata, Wilke & Czienskowski, 2009; Wilke, Gigerenzer & Jacobs, 2006). This way, we obtained a slope and intercept for each participant, where the intercept represented the initial time spent in the current display without a reward detection, and the slope represented the increase in the residence time with each new reward capture.

Individual slopes in both subgroups of human participants were in all cases positive and on average significantly above zero, [short-GUT: mean slope = ﻿0.235 ± 0.072 , *t*(20) = ﻿14.643, *p* < .001, long-GUT: mean slope = ﻿0.156 ± 0.050, *t*(20)= ﻿13.965, *p* < .001] (see Figure 8 a)). This suggested that participants indeed extended their residence time in response to a new reward capture, consistent with the incremental patch-leaving rule. Intriguingly, within-subject slopes in the short-GUT group of humans, that had shown less evidence for a GUT rule, were significantly steeper compared to the slopes of the long-GUT group, *t*(40) = 4.179 *p* < .001, indicating that the incremental effect of reward captures on the likelihood to stay in the current patch was stronger in short-GUT humans. We conducted the same within-subject regressions also for the short- and long-GUT group of gerbils and obtained a similar pattern of results. The mean slope was 0.0790 ± 0.048 in the short-GUT group, *t*(8) = ﻿4.154, *p* = .003, and ﻿0.095 ± 0.061 in the long-GUT group, *t*(8)=﻿4.394, *p* = .002. The average slopes did not differ between the two groups, *t*(16) = -﻿0.934, *p* = ﻿.364 (Figure 8 b)).

Taken together, in support of the incremental rule, within-subject regressions in both species showed that individuals extended their residence times incrementally following a new reward capture. Intriguingly, this incremental effect of reward capture on residence time was stronger in humans in the short-GUT group that had shown less evidence for a consistent GUT rule compared to the human long-GUT group. The same trend was observed in gerbils, yet lacking statistical significance.

Ein Bild, das Reihe, Screenshot, Diagramm, Steigung enthält.

Automatisch generierte Beschreibung

**Fig. 8. Incremental relationship between reward captures and residence times**. a) Regression plot on group level for all humans. Black dots show individual data points. The blue line shows the averaged regression line of the individual intercepts and slopes obtained from the within-subject regressions regressing residence times on the number of rewards for short-GUT humans. The green line shows the same for the long-GUT humans. The mean individual slope was significantly higher in the short- compared to the long-GUT group, indicating a stronger incremental relationship between reward captures and residence times in short-GUT humans. b) Plotted is the same data for the short- (yellow) and long-GUT gerbils (red). Here, individual slopes did not differ significantly between the two groups.

**Instantaneous and average collection rates - early gerbils, belated humans**

According to the marginal value theorem, optimal patch-leaving decisions are timed to the moment when the ICR approximates the MCR. As an estimate of the collection rate at which a reward capture *i* occurred, we used the inverse of the time that had passed between the previous reward capture *i-1* and the reward capture *i* (i.e., 1/intertarget times). To approximate the collection rate in the moment of patch-leaving, we used the inverse of the observed GUT, i.e., the time since the last reward capture and leaving the current patch (McNair, 1982). The MCR we obtained by dividing the total number of reward captures by the total search time (see Wolfe, 2013). Given the difference in the number of total earnings between short- and long-GUT human subjects, we predicted that the latter group of participants had extended their residence times longer than what is considered optimal according to the MVT (i.e., estimated ICRs at the time of leaving should be significantly lower than the average collection rate).

Figure 9 a) shows the trajectory of the ICRs as a function of target captures for the long-GUT human subjects and b) for the short-GUT group. Testing the MVT prediction, a one-way repeated measure ANOVA with type of time interval (ICRs for the three patch qualities as level 1-3, and the average collection rate as the level 4) yielded a significant main effect for the type of interval, *F*(3, 60) = ﻿102.502, *p* < .001. In line with our prediction, post-hoc contrasts showed that all three estimated collection rates at the time of patch-leaving were below the average collection rate of 0.165 ± 0.02 rewards/s, with *p* < .001 in all three patch types. Also in the short-GUT humans we found a significant main effect of the type of time interval , Friedman *F*(﻿2.905, 58.095) = ﻿3.457, *p* =.﻿023. However, post-hoc tests revealed that in high-quality patches, estimated ICRs at the time of patch-leaving were still significantly *above* the average collection rate, *p* = ﻿.015, no evidence for a difference was found in medium- and low-quality patches, with *p* = .277, and *p* = ﻿.614. Thus, at least in medium and low-quality patches we found evidence for an optimal timing of the patch-leaving according to the MVT.

Repeating the same analysis for the gerbil data, the one-way repeated measure ANOVA with the type of time interval as the single factor yielded a significant main effect in both groups of gerbils, [long-GUT: Friedman *F*(﻿2.778, 22.222) = 8.701, *p* < .001; short-GUT: *F*(﻿2.778, 22.222 = ﻿19.0, *p* < .001]. Post-hoc contrasts for the long-GUT gerbils showed that in low-quality patches the estimated collection rates at the time of leaving were still significantly higher than the MCR, *p* = .001. No evidence for such differences were found for medium- and high-quality patches, *p* = .185, *p* = .670. In short-GUT gerbils, ICRs in low- and high-quality patches were still significantly above the MCR, *p* = .001, *p* = .010, in medium-quality patches the evidence was anecdotal, *p* = .052. Figure 9 c) and d) show the trajectory of the ICRs as a function of target captures for the long-GUT and the short-GUT group of gerbils, respectively.

Ein Bild, das Reihe, Diagramm, Screenshot, parallel enthält.

Automatisch generierte Beschreibung

**Fig. 9**: **Collection rates.** Point plots show the ICRs for seven last reward captures averaged for the long-GUT (**a)**) and short-GUT humans (**b)**). Error bars equal ± standard error. The gray dashed line marks the averaged overall collection rate given by the number of total rewards divided by the total search time in seconds (shaded area indicates the 95% confidence interval). The estimated capture rate at the time of leaving was defined as the inverse of the participant’s GUT with short GUTs leading to higher estimated ICRs at the time of leaving compared to long GUTs. Point plots show ICRs for the long-GUT (**c)**) and short-GUT (**d)**) gerbils of the last three reward captures. Note that only the last three reward captures are plotted here because this was the average number of rewards obtained. In contrast to the MVT prediction, ICRs in gerbils were well above the MCR (red dashed line) at the time of leaving (i.e., the ICR at which the last reward was captured).

**Formal model testing of cues used to inform patch-leaving decisions**

Lastly, we used cox regressions to test different predictors that potentially increased or decreased the likelihood to stay in the current patch (e.g., Hutchinson et al., 2008). For this purpose, we used cox-regressions that model the impact of different factors on the probability to leave the current patch (i.e., the hazard ratio of patch leaving). The Cox proportional hazard model is a regression model typically used in epidemiology to find out the relationship between the survival time of a patient and one or more predictor variables (Bender, 2009). The model has also become widely used in the foraging literature (e.g., Hutchinson et al., 2008; Lottem et al., 2018) to model the residence times using the following hazard function:

h(t) = h0(t) \* exp(b1x1 + b2x2 + … + bnxn),

where t is the residence time, and h(t) the hazard function of the residence time, b indicates the impact of the predictor x on the probability to reside in the current patch. The resulting value exp(bi) is called the hazard ratio (HR) for the predictor *i*. It refers to the relative ‘risk’ of leaving the current patch for different levels of the i-th predictor. In other words, it quantifies the change in risk of leaving the current patch associated with a unit change in the *i*-th predictor. A HR greater than 1 indicates an increased risk of leaving the patch, while a hazard ratio less than 1 indicates a decreased risk. A HR of 1 indicates no change in risk.

The within-subject regressions with the number of reward captures as the predictor and the residence times as the outcome had already indicated an incremental relationship between the two variables, consistent with an incremental mechanism based on reward encounters driving patch-leaving decisions in both species. Given this finding, we entered the number of reward captures as the first predictor to the model and expected a HR below 1, i.e., a protective effect of the number of rewards decreasing the risk of patch-leaving. Since all humans and gerbils had positive slopes in the within-subject regressions, we expected this protective effect to be significant in both short- and long-GUT humans and animals. In addition, we used the averaged inter-target times between the last and the second before last reward capture, and between the second-last and the third last reward captures. This value could provide subjects with a good estimate of the recent collection rate. Increases in value should increase the risk of leaving, at least in the short-GUT group of humans that had shown collection rate data most closely in accordance with the MVT (see also Hutchinson et al., 2008). Thus, we expected a HR significantly larger than 1 in the short- but not long-GUT group of humans, indicating that only the former group of subjects used their recent reward capture rates as a cue for patch leaving.

In line with our prediction, the cox regression for the short-GUT humans revealed that, while controlling for the number of rewards, the hazard ratio (HR) for the averaged ICI was 5.01, 95% CI [1.30, 19.42], which means a substantial increase in the ‘risk’ of patch-leaving, Wald χ2(1) = ﻿2.33, *p* = .02. In other words, if the average ICIs of the last two target captures increased by 1 s, subjects were 5 times more likely to leave the current patch compared to no increase in the averaged ICI. Thus, consistent with the MVT, the cox model confirmed that the short-GUT humans were sensitive to declines in their current collection rates and timed their patch-leaving accordingly. Again in line with our prediction, the number of reward captures conversely appeared to be a *protective* factor regarding the ‘risk’ of patch-leaving as they decreased the risk to leave by 80% with each new reward capture, HR = 0.20, 95% CI [0.09, 0.44], Wald χ2(1) = ﻿-3.95, *p* < .005, consistent with the incremental rule.

Next, we computed the same cox regression model also for the long-GUT group. Again, we could confirm the results of within-subject regressions that had indicated an incremental relationship between reward captures and residence times in that the number of reward captures had a protective effect on the ‘risk’ of patch-leaving, decreasing the risk by 73%, HR = 0.27, 95% CI [0.13, 0.55], Wald χ2(1) = -3.59, *p* < .005. Intriguingly, in the long-GUT group, the average ICI had no effect on the ‘risk’ of patch-leaving HR = ​​3.89 95% CI [0.38, 39.63], Wald χ2(1) = 1.15, *p* = .25. Thus, in the long-GUT group, subjects’ estimates of the recent collection rate did not impact their patch-leaving. These findings are consistent with results from the collection rate data showing that long-GUT humans’ ICRs were not in agreement with the MVT.

The data obtained from gerbils with long- and short-GUT were much more consistent between the two groups. We did not find any significant differences in the effect of the number of rewards and averaged ICIs on the risk of gerbils’ patch-leaving between the two groups. Hence, the cox regression was carried out for the entire group of gerbils (pooled long- and short-GUT animals together). The results matched those of the humans in the short-GUT group: if the averaged ICI increased by one second, gerbils were almost 3.5 times more likely to leave the current patch, HR = 3.46, 95% CI [1.03, ﻿11.69], Wald χ2(1) = 2.00, *p* = .05, while a new reward capture decreased the ‘risk’ of leaving by 94%, HR = 0.06, 95% CI [0, 0.90], Wald χ2(1) = -2.04, *p* = .02, confirming that gerbils relied on the reward encounters they experienced (incremental rule) but also on their current collection rates (MVT) in order to make patch-leaving decisions.

While it was guaranteed that the animals commenced each session with hunger and motivation, it is important to consider that this motivation might diminish with satiation. Such a decrease in motivation could lead to reduced task commitment, potentially resulting in a higher occurrence of task-unrelated behaviors. This, in turn, could notably affect behavioral measures like the average collection rate. To explore this hypothesis, we performed a split-half analysis within each session, comparing the behavioral parameters of gerbils between the two session segments and across the three levels of patch quality. These results indeed revealed behavioral changes consistent with a decrease in task motivation. While the number of nose pokes remained constant throughout an experimental session, increases in the inter-poke intervals, residence times, and travel times suggested an increase in the frequency or duration of task-unrelated behaviors such as grooming. However, crucial parameters describing the animals’ patch-leaving behavior suggested that animals continued to optimize the timing of their patch departures despite the fading task-motivation. Although the average collection rate declined as function of session-split and were GUTs increased in the second compared to first half of a session, the difference between ICRs and MCRs was on average significantly smaller in the second half. Importantly, these results do not challenge the conclusion that we made based on the results reported previously. The details of this analysis are reported in the supplementary material.

**Discussion**

This study had the goal to elaborate a specific probabilistic foraging paradigm for inter-species comparisons. To this end, we designed two separate foraging tasks that were tailored to suit each species. Although the tasks differed in their specific details, the underlying reward structure was intentionally made comparable. This allowed us to investigate both the similarities and dissimilarities in foraging behavior between animals and humans.

**Humans and gerbils both adapt to changing reward probabilities.**

Timing patch departures based on a fixed number of reward captures or based on a fixed amount of time only works well if the forager roams an environment that offers patches that do not differ greatly in quality (see Wilke, 2004). However, both the fixed N- as well as the fixed-T rule are not optimal if patches within an environment differ greatly in quality (Iwasa et al., 1981; Stephens & Krebs, 1987). Thus, given the randomly changing reward probabilities in our paradigms, the use of these rules would have been disadvantageous for our species. The findings that both residence times and reward captures increase with increasing patch quality, confirmed this assumption. However, humans showed high variability in individual residence times and GUTs. We therefore divided the sample of humans based on the median GUTs into long- and short-GUT individuals to evaluate whether different rules apply to these two subgroups of human participants. For better comparison we did the same also for the group of gerbils. Both subgroups of both species showed evidence for the incremental rule of patch-leaving (see Figure 1 c)) and all animal and human subjects were sensitive to sunk costs. However, only the short-GUT group of humans but all gerbils timed their patch-leaving optimally according to the MVT. Lastly, a cox-regression suggested that both species used similar cues to time their patch-leaving.

**Reward captures incrementally delay patch-leaving in both species**

Due to the probabilistic decay function and the additionally varying initial probabilities, predicting the reward probabilities for both humans and gerbils was exceedingly challenging.

To tackle single reward encounters provided a first means for our gerbils and humans to learn about the quality of the current patch and each new reward encounter could be perceived as an indication that the current patch may be of good quality and thus worthwhile to spend more time in it. Consistent with this notion, the within-subject regressions, where we regressed residence times on the number of reward captures, showed a positive slope in all individuals of *both* species, demonstrating that residence times were extended incrementally following a new reward capture. This led to significantly longer residence times in high compared to low quality patches in both species and is consistent with an incremental mechanism driven by reward encounters that incrementally increase residence times by postponing the patch-leaving (Iwasa, 1981). Further support of this notion was given by the results of the cox regression in both species showing that the number of reward captures had a significant protective effect on the ‘risk’ of patch-leaving (i.e., delaying patch-leaving). This result confirms previous findings in humans (e.g., Wilke et al., 2009) and agrees with the results in mice by Lottem and colleagues (2018), from whom we adopted the probabilistic foraging task for our gerbils. However, in contrast to our analysis, Lottem et al. (2018) analyzed mice nose-pokes and fitted these data with a proportional hazard model. The estimated hazard rate reflected the probability to leave a current patch as a function of nose-pokes that started at its minimum with the beginning of a trial and would increase with each unrewarded nose poke. Each rewarded nose poke, however, decreased the hazard of leaving prolonging residence times. While the mice collection rates at the time of leaving were in keeping with the MVT, the incremental model showed a significantly better fit compared to a MVT-based model fitting. Using the same foraging paradigm in gerbils, but analyzing reward captures and residence times instead of nose-pokes, we can replicate this finding for gerbils. By designing a task with a foraging environment that has a comparable reward structure, we can also show that human foragers similarly adopt an incremental mechanism for patch-leaving. This demonstrates that both species adapt their foraging strategy in a similar way when facing an environment of unknown and variable reward structure, where some patches offer more rewards than others.

**Overharvesting only in humans**

Although the incremental rule was evident in the data of all humans, the relationship between reward encounters and residence times was weaker (significantly shallower slopes) in the long- compared to the short-GUT humans. The former group of human participants, however, showed GUT data most consistent with a simple GUT rule: their GUTs appeared to be consistent across patch-types, and always exceeded the previous ICIs. Moreover, it showed, these subjects had adopted a rather detrimental GUT rule with excessively long GUTs resulting in *overharvesting* (i.e., exploiting a patch longer than what is considered optimal according to the MVT). The disadvantageous timing of their patch-leaving became evident in the significantly poorer overall performance measured in terms of total reward earnings compared to short-GUT humans. In the short-GUT group of humans, in contrast, only a minority of subjects demonstrated foraging behavior with GUT rule-conforming GUT-ICI patterns. Hence, the strongest evidence for a fixed GUT rule in humans was found in subjects who had the tendency to overharvest.

There was no evidence for such a difference in performance between the two subgroups of gerbils. Both subgroups of gerbils had consistent GUTs across patch-qualities, and GUTs exceeded the ICIs in all but one animal. These findings are well in agreement with the predictions of the simple GUT rule. In accordance, also previous studies in other foraging animals facing patches of unpredictably varying quality reported patch-leaving behavior that was in agreement with a simple GUT-rule (e.g., Redhead & Tyler, 1988; Ydenberg, 1984). Yet, in stark contrast to the long-GUT humans who also used a GUT rule, the gerbils appeared to keep their GUTs relatively short, thus avoiding overharvesting.

A mundane interpretation for the phenomenon of overharvesting being exclusive to humans might suggest that these individuals possibly did not grasp the task. However, efforts to gauge subjects' comprehension through post-briefing questions and training performance assessment did not confirm this hypothesis. Moreover, prior to the main experiment, all subjects were told that an exhaustive search strategy (i.e., trying to find all existing rewards per display) would be detrimental to the overall task performance. Studies in elderly foraging humans reported age-related increases in GUTs as an indication of an increased behavioral tendency to exploit (e.g., Mata et al., 2013). This suggests that exploration and exploitation as opposing behavioral tendencies together form a continuum and that individuals may differ in their position along this continuum due to age differences and other factors. For instance, a recent study in patients with opioid use disorder showed that interindividual variability in overharvesting (in both users and controls) was related to a poorer neuromelanin signal, and indirectly catecholaminergic function (i.e., dopamine), of the VTA (Raio et al., 2022). Neurotypical subjects may already differ in their tendency to either explore or exploit based on genetic variations in those genes that control, e.g., the formation of the catecholaminergic system. Still, the extent to which long-GUT subjects tended to overharvest is striking, given the disadvantage that arose from this strategy. Clearly, more research is needed to further examine potential reasons that could explain these interindividual differences in foraging behavior.

**Evidence for the sunk-cost effect in both species**

When monitoring the humans performing the task, it showed that they tended to extend their GUTs depending on the amount of time they had already spent in the current display. This was confirmed by the results of the within-subject regressions between giving-up times and residence times. Here, all but four humans had positive slopes which is consistent with the sunk-cost effect (also referred to as the ‘Concorde fallacy’; e.g., Arkes & Ayton, 1999). This is a cognitive bias manifested as the inclination of individuals to persist in allocating resources to an ongoing endeavor or decision that has already accumulated substantial costs (referred to as sunk costs), even when the prospect of success is low or the initial choice was mistaken. Since only additional costs and benefits incurred by continuing with a particular course of action (i.e., marginal), but not past costs (i.e., in our case the residence time that the subject had already invested in foraging in the current patch), should factor into the decision making, the effect is considered maladaptive (Navarro & Fantino, 2005). Interestingly, we observed a parallel pattern of outcomes among gerbils. The presence of positive individual slopes indicated that these animals were also attuned to their prior time investments. This sensitivity was evidenced by the tendency of gerbils to tolerate longer durations of unsuccessful search as their cumulative time in the current patch increased, before they left the patch. Contrary to the initial assumption that only humans succumb to this bias (Arkes & Ayton, 1999), mounting evidence points towards its presence in diverse species, including rodents like mice and rats (Redish et al., 2022; Sweis et al., 2018b; Wikenheiser, Stephens & Redish, 2013), as well as avians and other creatures (for an overview, see Mahalgães & White, 2016; Pattison, Zental & Watanabe, 2012; Watzek & Brosnan, 2020). Our findings align with this growing body of evidence, providing further support to the notion that the sunk cost phenomenon extends across a range of species.

**Humans who forage more bias-free confirm MVT predictions**

GUT-rule-using human participants who exhibited a tendency to overharvest (referred to as the 'long-GUT group') demonstrated extended patch residence times that exceeded optimal values according to the MVT (Charnov, 1976). The MVT theory posits that foragers decide when to leave a patch based on diminishing returns. Specifically, they should depart when the ICR – representing the current rate of rewards obtained – approximates the MCR of the entire environment. This timing ensures that the energy spent on acquiring additional rewards does not surpass the benefits gained and is thus considered as optimal.

Notably, the overharvesting long-GUT humans displayed a bias toward exploitation, estimating ICRs at the point of patch departure significantly below the average collection rate. In contrast, short-GUT humans, who tended to leave patches earlier, aligned with MVT predictions by estimating ICRs comparable to the average collection rate. Intriguingly, this latter group of subjects included participants of whom less than half exhibited ICI patterns conforming to the GUT-rule.

Our findings suggest that while the most compelling evidence for a maladaptive GUT rule was found in humans with above-median GUTs, those who left patches earlier in time (the short-GUT group) showed collection rate patterns more consistent with MVT principles than with the GUT rule. This implies that these two groups may have employed different cues to determine patch departure. In the case of the GUT rule, a subjective temporal threshold for unsuccessful searching is employed. In contrast, subjects adhering to MVT principles may have used recent ICRs as an estimate of their present intake rate and a cue for timing patch departure. This hypothesis is supported by our cox regression results in short-GUT subjects, where the average of the last two inter-target intervals – an effective estimate of the current ICR – emerged as a robust positive predictor for patch-leaving. Increasing this interval (reflecting a decrease in estimated ICR) heightened the 'risk' of leaving the patch. Conversely, within the long-GUT group, the same predictor (current ICR) yielded no statistical significance.

A recent foraging study used a visual search paradigm similar to our human task and reported that human participants foraged longer in a given patch than predicted by the MVT (Kristjánsson et al., 2020). Their subjects performed either a conjunction (i.e., targets were defined by a combination of two features as in our experiment) or a feature search task (i.e., targets were defined by a single feature), and the results indicated similar foraging behavior for both search types. Yet, in contrast to the conjunction search task used in the present study, the volunteers were allowed to switch between target types within a given patch. During these switches the ICR would drop well below the MCR, but the subjects stayed in the same display. This behavior is inconsistent with the MVT prediction. Additional investigations employing visual search paradigms, such as virtual berry picking experiments, have revealed deviations from the MVT's projections under specific circumstances. These deviations are notably pronounced when patch quality exhibits significant variability, and when visual information becomes impaired to the extent that foragers are unable to discern whether a target item offers a reward (Wolfe, 2013, experiments 5 and 6). This evidence suggests that rendering foraging tasks more complex, e.g., by allowing changes between search types (Kristjánsson et al., 2020) or introducing a high degree of reward variability (Wolfe, 2013), patch-leaving behavior appears to be no longer in accordance with the MVT. In humans we showed that inter-individual differences in displaying a behavioral bias to exploit affect whether patch-leaving conforms to the MVT or not. A stronger exploitation bias leads to patch-leaving decisions that are less in keeping with the MVT in humans.

Testing the MVT in gerbils, animals in the long-GUT groups showed collection rates more in keeping with MVT. Only in low-quality patches these animals left the current patch when ICRs were still significantly above the average rate. Short-GUT gerbils had a higher propensity for early patch-leaving and had ICRs at the time of leaving that were still significantly above the average rate in two out of three reward conditions. Thus, when comparing gerbils’ ICRs and MCRs, the animals appeared to time their patch-leaving slightly earlier than predicted by the MVT. Nevertheless, the cox-regression suggested that animals indeed used an estimate of their ICR to support their patch-leaving decisions. Specifically, decreases of the ICR increased the risk of patch-leaving consistent with the MVT. Prior studies in animals have provided mixed evidence either in support (e.g., Lottem at al., 2018) or against the MVT (e.g., [DiGiorgio](https://onlinelibrary.wiley.com/authored-by/ContribAuthorRaw/DiGiorgio/Andrea+L.), [Upton](https://onlinelibrary.wiley.com/authored-by/ContribAuthorRaw/Upton/Elizabeth+M.), [Susanto](https://onlinelibrary.wiley.com/authored-by/ContribAuthorRaw/Susanto/Tri+Wahyu) & [Knott](https://onlinelibrary.wiley.com/authored-by/ContribAuthorRaw/Knott/Cheryl+D.), 2021; [Grether](https://link.springer.com/article/10.1007/BF02547724#auth-Gregory_F_-Grether), [Palombit](https://link.springer.com/article/10.1007/BF02547724#auth-Ryne_A_-Palombit) & [Rodman](https://link.springer.com/article/10.1007/BF02547724#auth-Peter_S_-Rodman), 1992). Our data suggests that although the gerbils’ ICRs were not perfectly aligned with the environment’s MCR at the time of patch-leaving, they were still sensitive to changes in their ICRs. The cox regression showed that decreasing ICRs increased the probability of patch-leaving to occur which is in accordance with the MVT.

**Differences in the reward regimes between gerbils and humans**

Comparing foraging behavior between the two species, a potential confound could have arisen from the differences in how the reward probability decreased in the two paradigms. In the gerbil task, the reward depleted with each executed nose-poke. In contrast, in the human task, the depletion followed a new reward capture but not a target fixation which would have been analogous to a nose-poke. Clearly, the reward decay following nose-pokes led to a faster depletion of the remaining reward probability. This could explain why the gerbils left patches more readily leading to very short residence times and low number of reward captures per patch compared to humans. Yet, despite these differences, there were commonalities between the species. Especially, both species consistently demonstrated an incremental mechanism driving patch-leaving. This finding agrees with that in a previous report (Lottem et al., 2018). Similarly, both species, except for the overharvesting human subjects, timed their patch-leaving in accordance with the MVT. Importantly, the unpredictable variability of reward across patches was the same in both experiments, thus, both species had to adapt to similar environments with aggregated reward distributions. Therefore, despite the difference in how the reward decayed we are confident that the two tasks were still sufficiently similar to comparatively study patch-leaving decisions of humans and rodents. Nevertheless, future investigations should eliminate the difference in the reward decay as potential confound to further improve the alignment of the two tasks.

**Outlook**

To better understand the neural coding mechanisms underlying patch-leaving decisions, it will be crucial to use fMRI in humans and recordings of local field potentials in gerbils. A central structure enabling attentional exploration in humans is the anterior prefrontal cortex, mainly consisting of Brodmann area 10 (aPFC; (Pollmann, 2016; Mansouri et al., 2017; Zajkowski et al., 2017)). Lesions of the aPFC, both in human and non-human primates (Pollmann et al., 2007; Mansouri et al., 2015) prevent the exploratory allocation of attentional resources to novel aspects of the environment, thereby preventing optimal adaptation to the environment. Although aPFC has a causal role for exploratory attention shifts in primates, rodents, lacking a distinct aPFC (Laubach et al., 2018), undeniably show exploratory behavior (Birke et al., 1985). This leads to the question how the rodent brain supports exploratory and rule-based (conditioned) attentional resource shifts. Rodent experiments are therefore not only of interest from a phylogenetic perspective, but also open broad opportunities of mesoscale investigations. Moreover, as the behavioral data reported here suggest that humans display strong interindividual differences in their tendency to exploit a current patch, future studies should target the question whether such behavioral differences also translate to physiological differences, such differential activations of the FPC and other areas of the frontoparietal attention network.

**Conclusions**

The present findings emphasize the value of inter-species comparisons in assessing the alignment of ecological models of animal behavior with the decision-making processes of humans during foraging. Our results indicate that rodents and humans can embrace similar decision rules within environments with comparable levels of reward unpredictability. Concurrently, the results highlight the substantial variability inherent in human decision-making during foraging. Moreover, these individual distinctions might be influential in dictating the extent to which human behavior adheres to the ecological models characterizing animal behavior.

**Authors’ contributions**

S.P. and M.H. conceived the idea and acquired funding for the project. S.P. and L.G. designed the visual search task for the human participants. L.G. realized humans’ data collection supported by student assistants, wrote all python scripts for data analysis, conducted the analysis of the human and the animal data, and drafted the original manuscript. F.O., M.H., and M.V. conceived and supervised the animal experiments. P.S. performed animal training and data collection. All authors commented on the final version of the manuscript.

**Acknowledgements**

The study was funded by the Deutsche Forschungsgemeinschaft, Project-ID, to SP and MH, project-ID 425899996e CRC 1436 (sub-project C02). We would like to thank Gabriel Dobroschke and Sara Claassen for their help with human participants’ data collection.

**Code availability**

The original costume python code used for data analyses can be accessed from <https://osf.io/fexgb/> as well as on git-hub: https://github.com/LGparrot/exploratory-attention-in-visual-foraging.

**Data availability**

Both animal and human data can be accessed from <https://osf.io/fexgb/> .

**References**

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals?. Psychological bulletin, 125(5), 591. doi: https://doi.org/10.1037/0033-2909.125.5.591

Bender, R. (2009). Introduction to the use of regression models in epidemiology. Cancer Epidemiology, 179-195.

Birke, L. I., D'Udine, B., & Albonetti, M. E. (1985). Exploratory behavior of two species of murid rodents, Acomys cahirinus and Mus musculus: a comparative study. *Behavioral and Neural Biology*, *43*(2), 143-161.

Cain, M.S., Vul, E., Clark, K., Mitroff, S.R. (2012) A Bayesian optimal foraging model of human visual search. *Psychological Science 23*(9): 1047–1054.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, *9*(2), 129-136.

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. Nature, 441(7095), 876-879. https://doi.org/10.1038/nature04766

Fujino, J., Fujimoto, S., Kodaka, F. *et al.* (2016). Neural mechanisms and personality correlates of the sunk cost effect. *Scientific Reports* 6, 33171. <https://doi.org/10.1038/srep33171>

Laureiro‐Martínez, D., Brusoni, S., Canessa, N., & Zollo, M. (2015). Understanding the exploration–exploitation dilemma: An fMRI study of attention control and decision‐making performance. Strategic management journal, 36(3), 319-338. doi: https://doi.org/10.1002/smj.2221

Horowitz T.S., Thornton I.M. (2008). Objects or locations in vision for action? Evidence from the MILO task. *Visual Cognition* *16*(4):486–513.

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

Krebs, J. R., Ryan, J. C., & Charnov, E. L. (1974). Hunting by expectation or optimal foraging? A study of patch use by chickadees. *Animal Behavior*, *22*,953–964.

Kristjánsson, Á., Björnsson, A. S., & Kristjánsson, T. (2020). Foraging with Anne Treisman: Features versus conjunctions, patch leaving and memory for foraged locations. *Attention, Perception, & Psychophysics*, *82*(2), 818-831.

Lottem, E., Banerjee, D., Vertechi, P., Sarra, D., Lohuis, M.O., Mainen, Z.F. (2018). Activation of serotonin neurons promotes active persistence in a probabilistic foraging task. *Nature Communicatio*n, *9*:1–12.

Mansouri, F. A., Buckley, M. J., Mahboubi, M., & Tanaka, K. (2015). Behavioral consequences of selective damage to frontal pole and posterior cingulate cortices. *Proceedings of the National Academy of Sciences*, *112*(29), E3940-E3949.

Mansouri, F. A., Koechlin, E., Rosa, M. G., & Buckley, M. J. (2017). Managing competing goals—a key role for the frontopolar cortex. *Nature Reviews Neuroscience*, *18*(11), 645-657.

Navarro, A. D., & Fantino, E. (2005). The sunk cost effect in pigeons and humans. Journal of the experimental analysis of behavior, 83(1), 1-13.

Pattison, K.F., Zentall, T.R., Watanabe, S. (2012). Sunk cost: pigeons (Columba livia), too, show bias to complete a task rather than shift to another. Journal of Comparative Psychology, *126*(1):1-9. doi: 10.1037/a0023826. Epub 2011 May 16. PMID: 21574688.

Pollmann, S., Mahn, K., Reimann, B., Weidner, R., Tittgemeyer, M., Preul, C., ... & von Cramon, D. Y. (2007). Selective visual dimension weighting deficit after left lateral frontopolar lesions. *Journal of Cognitive Neuroscience*, *19*(3), 365-375.

Redhead, E., & Tyler, P. A. (1988). An experimental analysis of optimal foraging behaviour in patchy environments. *The Quarterly Journal of Experimental Psychology Section B*, *40*(1b), 83-102.

Sih, A., (2013). Understanding variation in behavioural responses to human-induced rapid environmental change: a conceptual overview. *Animal Behavior*, 85:1077–1088.

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald, A. W., Thomas, M. J., et al. (2018a). Sensitivity to “sunk costs” in mice, rats, and humans. *Science* 361, 178–181.

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software*, 3(31), 1026.<https://doi.org/10.21105/joss.01026>

Wajnberg, E., Bernhard, P., Hamelin, F., & Boivin, G. (2006). Optimal patch time allocation for time-limited foragers. *Behavioral Ecology and Sociobiology*, *60*(1), 1-10.

Watzek, J., Brosnan, S.F., (2020). Capuchin and rhesus monkeys show sunk cost effects in a psychomotor task. *Scientific Reports* 10, 20396. https://doi.org/10.1038/s41598-020-77301-w

Wilke, A., Hutchinson, J. M., 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.

Wolfe, J.M. (2013) When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision,* 13(3):1–17.

Zajkowski, W. K., Kossut, M., & Wilson, R. C. (2017). A causal role for right frontopolar cortex in directed, but not random, exploration. *Elife*, 6, e27430.