

Published in final edited form as:

Nat Hum Behav. 2019 July ; 3(7): 733–745. doi:10.1038/s41562-019-0603-9.

Minimizing threat via heuristic and optimal policies recruits hippocampus and medial prefrontal cortex

Christoph W. Korn^{1,2,3} and Dominik R. Bach^{1,2,4}

¹Computational Psychiatry Research, Department of Psychiatry, Psychotherapy, and Psychosomatics; Psychiatric Hospital, University of Zurich, Lengstrasse 31, 8032 Zurich, Switzerland ²Neuroscience Center Zurich, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland ³Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany ⁴Wellcome Trust Centre for Neuroimaging and Max-Planck UCL Centre for Computational Psychiatry & Ageing, University College London, 12 Queen Square, London WC1N 3BG, United Kingdom

Abstract

Jointly minimizing multiple threats over extended time horizons enhances survival. Consequently, many tests of approach-avoidance conflict incorporate multiple threats for probing corollaries of animal and human anxiety. To facilitate computations necessary for threat minimization, the human brain may concurrently harness multiple decision policies and associated neural controllers—but it is unclear which. We combine a task that mimics foraging under predation with behavioural modelling and functional neuroimaging. Human choices rely on immediate predator probability—a myopic heuristic policy—and on the optimal policy, which integrates all relevant variables. Predator probability relates positively and the associated choice uncertainty relates negatively to activations in anterior hippocampus, amygdala, and dorsolateral prefrontal cortex. The optimal policy is positively associated with dorsomedial prefrontal cortex activity. We thus provide a decision-theoretic outlook on the role of the human hippocampus, amygdala, and prefrontal cortex in resolving approach-avoidance conflicts relevant for anxiety and integral for survival.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence: Christoph W. Korn, Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany, c.korn@uke.de.

Code availability

The code used for the analyses is available at github (https://github.com/dnhi-lab/minimizing_threat.git).

Data availability

The behavioural data that support the findings of this study are publicly available at github (https://github.com/dnhi-lab/minimizing_threat.git) and at figshare (<https://doi.org/10.6084/m9.figshare.7929914.v1>). The neuroimaging data that support the findings of this study are publicly available at neurovault (<https://neurovault.org/collections/5046/>).

Author contributions

C.W.K and D.R.B designed the experiment, developed the analysis procedures, and wrote the paper. C.W.K. collected and analysed the data.

Competing interests

The authors declare no competing interests.

Christoph W. Korn: <http://orcid.org/0000-0002-8228-2624>

Introduction

In order to survive animals must minimize threats such as dying from starvation or predation. This often entails conflicts between approaching food and avoiding predators. Thus, decisions fundamental for survival necessitate challenging calculations to balance competing goals, such as simultaneously maintaining energy homeostasis and physical integrity 1. In principle, this can be achieved in a normative manner by jointly minimizing the expected impact of different threats such as starvation and predation. To do so, decision-makers should consider extended sequences of risky decisions, in which threat occurrences depend probabilistically on both current and future decisions 2,3. But computing the optimal policy for such multi-step decision situations requires taxing evaluations of all possible future states, which might be too complicated to execute under threat 4–7. Therefore, decision-makers may take advantage of simplifying, heuristic policies that approximate optimal solutions, for example by minimizing the most prominent immediate threat and disregarding future time points. Consequently, it has been proposed that the human brain comprises multiple neural controllers which implement a variety of decision policies spanning different levels of sophistication and efficacy 1,6,7. Several of these policies and neural controllers may be concurrently invoked when seeking a solution to a complex situation that threatens survival 2,8–10.

A particularly relevant example for such decision problems is foraging for food under the risk of predation, which involves an approach-avoidance conflict (i.e., approaching food while avoiding predators). Laboratory tasks that mimic this scenario are widely used as animal anxiety tests, in which an animal can explore the environment or obtain food amid simultaneous physical threat 11–16. Some of these paradigms have been successfully reverse-translated to human computer games 17–25. Importantly, a long-standing tradition of lesion studies in rodents 16 and an emerging field of lesion and functional imaging studies in humans 17–22,24,26,27 converge on implicating similar neural structures—in particular the anterior hippocampus and the amygdala. These findings suggest an at least partly homologous neural implementation of approach-avoidance conflicts.

Despite progress in understanding the neural circuits required for decision-making in approach-avoidance conflicts 10,28,29, the algorithms used in such scenarios, and their neural implementation—especially with respect to the hippocampus—remain elusive. Most paradigms do not separate individual decisions (such as the elevated plus maze, the open field test, and their human analogues) and several others reduce the task to a single decision with relatively low complexity (such as operant conflict tests in rodents and humans). In these tasks, it is thus unclear what goal the agent pursues and whether momentary behaviour is part of an extended action plan. Furthermore, previous research has rarely assessed the agent's evaluation of different risky outcomes in approach-avoidance conflicts (see the study on macaques by 30 for a notable exception). This leads to a disconnection between neurobiological research on approach-avoidance conflicts, which has implicated the hippocampus, and decision neuroscience, which has mostly linked (medial) prefrontal regions to flexible action selection 31–36.

Here, we apply a decision-theoretic outlook on choice sequences in approach-avoidance conflict by employing a mathematically specified computer game that mimics foraging under predation. We first address the cognitive strategies by which human decision-makers resolve approach-avoidance conflicts, and then investigate the neural representation of the associated decision variables (DV), specifically of a threat-related DV in hippocampus and the amygdala. This neural hypothesis elaborates our behavioural hypothesis that humans base their decisions primarily on a threat-related policy. The design of our task was inspired by several recent studies using virtual foraging tasks in humans and non-human primates 2,3,37–42 and thus links approach-avoidance conflicts to the burgeoning literature on complex decision behaviour in ecological scenarios 1,8–10,43.

Results

An approach-avoidance task afforded the computation of an optimal policy

Our approach-avoidance conflict task was framed as virtual foraging under the dual threats of predation and starvation (see Figure 1 for a task outline and Table 1 for a list explaining all relevant variables). In brief, participants made sequential decisions on up to five trials, called “days,” in 240 mini-blocks, called “forests” (resulting in total of 400 trials, i.e., days, per participant). Participants were monetarily incentivized to (a) keep their “energy points” above zero across the possible maximum number of five days within a given forest and to (b) attain as often as possible the maximum number of five energy points. Otherwise, the exact number of energy points accrued in a forest was not translated into financial payoffs. On each day within a forest, participants had to make a decision between a certain option, called “waiting,” and a risky option, labelled “hunting” which we refer to here as “foraging.” Waiting entailed a sure loss of one energy point. Foraging had several possible outcomes: first, participants could be “attacked by a predator” with known probability, and lose all their energy points on the current day. If they were not attacked, the known “probability of foraging success” determined whether they would gain a variable number of energy points, or lose a fixed number of two energy points. Thus, participants could reach zero energy points due to “predation” or due to “starvation.” On each day within a forest, one of two combinations of probabilities of predator attack and foraging success was randomly presented, called “weather type.” Conditions, in terms of probabilities of predator attack and foraging success, appeared in randomized order. Participants were fully and explicitly informed about all variables of the current forest. They knew that they would stay within each forest for a maximum of five days, and that they would be rewarded for a full five-day sequence. The number of days past was not depicted on the screen.

The task was mathematically specified as a Markov decision process. This allowed us to calculate the *a priori* optimal policy that maximizes participants’ monetary rewards by jointly minimizing the threats of dying due to predation and starvation for a fixed time horizon of five days (as well as maximizing the number of times reaching the maximum energy level). This optimal policy combines the probabilities of predator attack and foraging success as well as further task variables in a mostly non-linear fashion (Supplementary Figure 1). We use the word “optimal” here in the sense of “optimal under task instructions” (i.e., an optimal choice implies choosing the action that maximizes the weighted sum across

the relevant Markov branches). The optimal policy *per se* makes prescriptions for choice on the basis of the value difference between the two choice options of foraging and waiting. That is, the optimal policy *per se* prescribes foraging if the value difference exceeds zero, it prescribes waiting if the value difference falls below zero, and it is indifferent if the difference is exactly zero. We related participants' choices, reaction times (RTs), and functional magnetic resonance imaging (fMRI) data to the continuous range of value differences according to the optimal policy. In the following, we therefore use the term "optimal policy" as shorthand to refer to this continuous range of value differences between foraging and waiting.

A large set of decision variables could potentially explain participants' behaviour

To explain participants' choices, we considered an extensive set of potential DVs: (1) the optimal policy and (2-16) fifteen possible DVs that are not optimal and are therefore considered as heuristics. In general, we considered heuristics for three reasons: because they constituted components of our sequential decision-making task, because they were imperfect variants of the optimal policy, or because they captured influences of past trials (see Table 1 for detailed descriptions).

Specifically, three heuristic were features of the current foraging option: (2) the predator probability, (3) the probability of foraging gain, and (4) the gain magnitude. Four variables took the current energy state into account: the current energy state as a (5) continuous variable ranging from one to five and as a (6) binary variable distinguishing the energy state one, in which waiting leads to sure death, from higher energy states. The two variables called (7) "expected energy" and (8) "expected energy change" correspond to the expected value of foraging at the current energy state, and the resulting difference in energy state. Another DV relied on (9) the number of days past. We refer to the following three heuristics as "pseudo-optimal policies" because they would be optimal in alternative scenarios: (10) one short-sighted policy would be optimal if participants remained in a forest for the current day only (i.e., time horizon of one day); two pseudo-optimal policies would be optimal if participants had to solely minimize the threats of either (11) starvation of (12) predation. Four further DVs were included to address suggestions by anonymous reviewers: (13) one of these DVs was a pseudo-optimal policy that weighted the time horizon of the optimal policy exactly according to the distribution of the number of days that participants actually remained in a forest (i.e., an average of 2.5 days, which was implemented to enhance fMRI design efficiency). Three DVs addressed the influence of the preceding trial or forest on the current trial: (14) the energy change from the last to the current trial, (15) "win-stay-lose-shift," and (16) a DV indicating whether participants had died in the preceding forest.

Participants' choices relied on predator probability and optimal policy

We then greedily searched for a single DV that best explained participant's choices. Bayesian model comparison identified predator probability as the best single predictor of participants' decisions out of the 16 candidate variables in the fMRI sample (final $n=24$; see Figure 2a for metrics of Bayesian Information Criterion (BIC) and Supplementary Table 1 for BIC and protected exceedance probabilities). Searching for the predictor that best explained the remaining variance revealed that participants additionally relied on the optimal

policy (Figure 2b). The model including both predator probability and the optimal policy outperformed the simpler model that only included predator probability and also more complex models that additionally included the interactions between different task variables (Supplementary Table 1), as well as all 66 possible combinations of two DVs from the set of the twelve DVs motivated *a priori* (the additional four DVs suggested in the review process were not included in this comparison because they had already performed quite badly in the initial analysis; Supplementary Table 2).

Crucially, the winning model robustly predicted empirical choices. This is illustrated in posterior predictive checks splitting data according to the two variables incorporated in the model (predator probability, Figure 2c, optimal policy, Figure 2d) and in posterior predictive checks splitting data according to the variables not included in the model (Supplementary Figure 2). Conversely, models with the other DVs did not capture the pattern of empirically observed choice (Supplementary Figure 3). Individual differences in the best fitting models are illustrated in Supplementary Figure 4.

We did not find evidence that the best-fitting DVs changed over the time course of the experiment. Specifically, we fitted models separately to data from the first and second halves of the experiment. In both halves, a combination of predator probability and optimal policy emerged as the best model (Supplementary Table 3).

All of the above results were replicated in an independent behavioural sample acquired during the revision process ($n=23$; see Figure 2 and Supplementary Tables 4-6). That is, predator probability decisively emerged as the single best predictor. Results according to BIC, which was our primary analysis approach, favoured a model combining predator probability and optimal policy (protected exceedance probability favoured a combination of predator probability and “binary energy;” Supplementary Table 4).

The two samples only differed when searching for a third DV that might explain variance on top of predator probability and optimal policy. In the fMRI sample, the comparison of models with three DVs suggested that the probability of foraging gain might best capture remaining variance (Supplementary Table 2). In the behavioural sample, “binary energy” emerged as the DV capturing most remaining variance (Supplementary Table 4). Due to the difference between the two samples we refrain from drawing conclusions about the identity of a potential third DV.

Behavioural models were distinguishable

The large set of DVs tested here entailed that some of these DVs were related to each other by design. For example, the optimal policy shared on average variance $> 50\%$ with two of the pseudo-optimal policies and with “expected energy change” (see Supplementary Table 7 for mean shared variances). Predator probability shared variance $> 50\%$ with two of the pseudo-optimal policies and with “expected energy.” Still, we argue that the DVs that emerged as relevant in our comparison were reasonably dissociable or explained variance on top of one another, as was the case for the optimal policy and predator probability. To illustrate the features of the optimal policy, we plotted the relations between the 15 heuristics and the optimal policy (Supplementary Figure 1).

More importantly, confusion analyses with 2000 simulations per model showed that models with one of the 16 considered DVs could be almost perfectly recovered from simulated data. This means it is unlikely that our winning model would have wrongly been selected if another model were the true model. Among the first 13 DVs of our list, only one simulated model was misclassified (Supplementary Table 8). Less than 10% misclassifications occurred for the last three DVs of our list that captured the influence of preceding trials or forests and that performed overall worst in the model comparisons (Supplementary Table 8). Confusion analyses on the models with predator probability plus one of the 15 other considered DVs indicated very good recovery of the respective models (Supplementary Table 9). Less than 5% misclassifications emerged for the initially considered DVs (but again misclassifications occurred more often for the last three DVs in our list that did not explain participants' behaviour; Supplementary Table 9).

To explore how participants' behaviour in the task was related to their subjective assessments, we administered a post-experiment questionnaire that assessed how much participants relied on different components of the task (Supplementary Table 10). Numerically, importance ratings were highest for predator probability, which may suggest that heuristic identified in the model comparisons corresponded to participants' consciously accessible decisions.

Reaction times scaled with optimal policy and choice uncertainty of predator probability

We predicted that participants' use of the predator probability and of the optimal policy should be reflected in their RTs. Specifically, the two metrics themselves and/or their corresponding choice uncertainties should be associated with RTs. This was indeed the case as shown by analyses of RTs in linear mixed effects (see Table 2 for statistics of the fMRI sample and the behavioural sample). In both samples, the optimal policy was directly related to RTs such that higher values for foraging versus waiting were related to faster decisions (fMRI sample: $n=24$, $t(30.67)=-3.04$, $p=0.005$; log-likelihood difference, $LLD=-4.0$, 95% confidence interval, $CI=[-0.10, -0.02]$; behavioural sample: $n=23$, $t(35.64)=-3.03$, $p=0.005$, $LLD=-4.1$, $CI=[-0.09, -0.02]$). Conversely, higher choice uncertainty of the predator probability was associated with slower decisions (fMRI sample: $n=24$, $t(25.17)=3.43$, $p=0.002$, $LLD=-4.8$, $CI=[0.22, 0.87]$; behavioural sample: $n=23$, $t(28.31)=2.69$, $p=0.012$, $LLD=-3.2$, $CI=[0.07, 0.54]$). Additionally, the discrepancy in choice probabilities according to the two metrics scaled with longer RTs (fMRI sample: $n=24$, $t(28.85)=5.53$, $p<0.001$, $LLD=-10.1$, $CI=[0.13, 0.27]$; behavioural sample: $n=23$, $t(22.34)=3.56$, $p=0.002$, $LLD=-4.9$, $CI=[0.05, 0.20]$). This discrepancy quantifies how much the prescriptions for choice differ between using predator policy and optimal policy. In sum, RTs corroborate that the predator probability and the optimal policy exert a joint influence on participants' decision process on the same trials.

fMRI data could be related to decision variables derived from the winning model

We aimed at identifying trial-by-trial associations of the behaviourally relevant variables with fMRI data during the choice phase. Specifically, we included the following variables as parametric modulators of the choice phase in our primary general linear model (GLM): (1) the predator probability, (2) the DV under the optimal policy, (3&4) the associated choice

uncertainties of the former two metrics, (5) the discrepancy between these two metrics, and (6) log-transformed RTs. In our primary GLM, parametric modulators were not orthogonalized but we obtained the same results (except for a few minute differences due to rounding) in separate GLMs when varying the order of the respective orthogonalized parametric modulators such that the relevant parametric modulators were entered last (see Supplementary Table 11 for shared variances between the included variables).

FMRI analyses related hippocampus, amygdala, and DLPFC activity positively to predator probability and negatively to the associated choice uncertainty

In line with our expectations derived from studies on related types of human approach-avoidance conflicts tests 18,26,27, predator probability was positively related to BOLD signals in a cluster in the right anterior hippocampus extending into neighbouring amygdala, as well as bilateral clusters in the dorsolateral prefrontal cortex (DLPFC; Figure 3a and Table 3). Notably, a cluster in the anterior hippocampus and neighbouring amygdala showed stronger BOLD responses with lower choice uncertainty of the predator probability (Figure 3b). This cluster overlapped with the cluster identified for predator probability *per se* (see Figure 4a for the overlap of the two functional clusters and for information on the overlap with anatomical hippocampus and amygdala masks, see Figure 4b or relation of parameter estimates to predator probability). Uncertainty of the predator probability was also negatively associated with BOLD signal in bilateral DLPFC and in right lateral inferior frontal gyrus (IFG) as well as other regions. Similar to the pattern in the hippocampus and amygdala, activity maps showed overlaps between the clusters identified for predator probability *per se* and the choice uncertainty of the predator probability (Supplementary Figure 5).

Taken together, both the predator probability and its associated choice uncertainty scaled on a trial-by-trial basis with BOLD response in the anterior hippocampus and the amygdala as well as DLPFC regions.

FMRI analyses related DMPFC activity positively to both the optimal policy and the associated choice uncertainty

The optimal policy was positively related to BOLD signals in the posterior dorsomedial prefrontal cortex (DMPFC; Figure 3c and Table 3), extending into the supplementary motor area. The same parametric contrast revealed a positive relation of the optimal policy with anterior cingulate cortex, bilateral IFG (extending into insula), and bilateral thalamus. The choice uncertainty of the optimal policy scaled positively with BOLD signal in DMPFC and DLPFC regions as well as IFG and insula (Figure 3d). The same metric scaled negatively with activity in the posterior cingulate cortex. In the DMPFC and IFG (extending into insula), we found overlaps between clusters elicited by the optimal policy *per se* and by the choice uncertainty of the optimal policy (Supplementary Figure 5). Furthermore, BOLD signal in the thalamus showed a positive association with the discrepancy in choice probabilities between the predator probability and the optimal policy.

FMRI analyses related activity in striatum and medial regions to outcomes

During the time point when participants saw the outcome of their choice, the change in energy state (i.e., energy state at outcome minus energy state at choice) was robustly associated with neural activity in a well-established reward network (i.e., bilateral striatum, ventral medial prefrontal cortex, and posterior cingulate cortex; Supplementary Figure 6 and Supplementary Table 12). This activation pattern was expected since participants were monetarily rewarded for avoiding an energy state of zero and for reaching the maximal energy state of five points.

Follow-up and exploratory fMRI analyses strengthen and illustrate the relationships between BOLD signal and behavioural variables

To exclude that the above-mentioned fMRI results were unduly driven by participants' choices themselves, we ran a secondary model, in which we additionally included choices as a binary parametric modulator. The clusters obtained with this secondary model generally replicated the results reported above (see Supplementary Table 13).

In the same vein, we also obtained similar clusters in a tertiary GLM, in which the choice uncertainties and the discrepancy were calculated from the independent behavioural sample (see Supplementary Table 14; the main difference was that the cluster in the left hippocampus and amygdala, which was positively related to the predator probability, survived small-volume correction for an anatomical mask of bilateral hippocampus but failed to reach whole-brain FWE correction at $p < 0.05$).

To explore inter-individual variability, we included the parameter estimates from the behavioural models linking participants' choices to predator probability and optimal policy as covariates into the respective contrasts. Two clusters in bilateral striatum co-varied negatively with individual parameter estimates capturing the degree to which predator probability influenced participants' choices (Supplementary Figure 7 and Supplementary Table 15). No significant clusters emerged for individual parameter estimates related to the optimal policy.

As noted in the previous sections, we observed that in some regions activity related to predator probability or optimal policy overlapped with activity related to the choice uncertainty of that DV (Supplementary Figure 5). Following the suggestions of anonymous reviewers, we visualized these relationships using functionally defined regions of interest (ROIs) based on the analyses described for our primary GLM. These analyses are not independent and serve illustrative purposes. Visual inspection shows that in several ROIs the relationships were not completely linear across the four quartiles of the parametric modulators, which likely suggests that these regions were influenced by more than one behavioural variable (Supplementary Figures 8, 9). To explore this pattern further, we conducted non-independent and post-hoc tests within each of the functionally defined ROIs to assess the specificity of these ROIs for the respective parametric modulators (at $p < 0.001$). As could be expected from the overlap maps (Figure 4 and Supplementary Figure 5), the clusters in the hippocampus and the amygdala as well as the clusters in bilateral DLPFC were positively related to predator probability *per se* and negatively related to the associated

choice uncertainty (Supplementary Table 16). Similarly, DMPFC and right IFG (extending into insula) were positively related to both the optimal policy *per se* and the choice uncertainty of the optimal policy. We also tested in a similar post hoc fashion whether BOLD signals in the identified functional ROIs were related to activity elicited when including the three next-best DVs from the behavioural model comparison into separate GLMs (i.e., the probability of foraging gain, continuous energy, and expected energy change; see Table 1). Only clusters in the medial occipital cortex showed relationships (Supplementary Table 16). Furthermore, supplementary GLMs that included the interactions between each policy and its associated choice uncertainty revealed clusters in the occipital lobe for these interaction contrasts (Supplementary Table 17).

Discussion

We demonstrate that humans employ two decision policies of varying complexity in a virtual approach-avoidance conflict task, which translates the often-evoked biological example of foraging under predation 9,21 into a mathematical framework amenable to decision-theoretic analyses. Participants primarily based their choices on the probability of predator attack—a myopic but easy-to-compute heuristic policy. Beyond that, they relied on the normatively optimal policy, which entails sophisticated integration of various task components as indicated by analyses of choice and RT data. These two policies were reflected in macroscopically different brain regions, which corroborates the theoretical notion that multiple neural controllers take care of different survival-relevant threats 1,8. Crucially, our results identify the neural controller of the heuristic policy with structures often implicated in approach-avoidance conflicts in both rodents and humans 11–16,18,21. That is, the anterior hippocampus and the amygdala related to predator probability as well as to the uncertainty of using this policy during choice. The optimal policy, and also the choice uncertainty thereof, were associated with parts of the DMPFC, which dovetails with the general roles of this region in decision-making 33,36.

Predator probability emerged as the primary policy employed by participants among a variety of potential alternatives. This result resonates with previous demonstrations of the same metric modulating behaviour in approach-avoidance conflict tasks with spatial layouts 18,26,27. Notably, participants used the predator probability as a primary heuristic and not the probability of foraging gain, which constituted the winning policy in an analogous virtual foraging task that did not include predation 2. This may indicate that participants are able to select a particularly appropriate heuristic for the task at hand, i.e., a heuristic that combines computational simplicity with near-optimal approximations. Given several theoretical accounts from multiple fields arguing for adaptive heuristic decision-making 43–47, it is an interesting avenue for future research to investigate how participants select and switch between different heuristics under varying biologically inspired decision tasks 2,3.

Neurally, different heuristics are likely to be implemented by different controllers 1. Here, the heuristic of using predator probability related to the anterior hippocampus and the amygdala, which demonstrates that in humans these structures are implicated in computing a decision policy central for a rather abstract type of approach-avoidance conflict without physical threats in contrast to most rodent tasks 12–16,48. Our fMRI results cannot delineate

subtleties in the roles of hippocampus and amygdala but we would like to highlight that work on rodents has identified dense monosynaptic connections between the two structures as well as reciprocal electrophysiological interactions during approach-avoidance conflict 15. While these connections appear to be crucial for avoidance of acquired threat predictors (as in the present study), they may be less relevant in avoiding innate threat predictors such as in the elevated plus maze 49.

With respect to work on humans, our findings relating anterior hippocampus and also the amygdala to predator probability corroborate recent studies implicating these structures in different types of approach-avoidance conflict 17–22,24,26,27. In contrast to several of these studies, our task did not entail spatial layouts, directional movements, or mnemonic demands, which have been argued to possibly impact hippocampal activity 21. The relation between predator probability and hippocampus in our rather strategic task fits particularly well with the recently demonstrated involvement of the hippocampus involvement during strategic decisions to escape from a slow-attacking virtual predator 24. Overall, the findings linking the hippocampus and amygdala with risks metrics related to virtual predators may thus suggest a more generic role of these regions in risky decision-making than usually acknowledged in the field of decision neuroscience. Future studies will be necessary to delineate whether these two structures are specifically involved when risks are framed in terms of threats to virtual survival. Possibly, participants in our tasks might have interpreted predator probability as some form of “defensive distance” since we visually depicted increasing predator probability as triangles of increasing sizes. Thus, our work accords with previous studies on defensive distance in rodents and recent efforts to reverse-translate defensive patterns from rodents to humans 50.

Interestingly, we found that the anterior hippocampus and the amygdala, in addition to tracking predator probability, also related to the choice uncertainty associated with using predator probability as a heuristic. Previous research has linked activity in the hippocampus to metrics of outcome uncertainty when humans make value-based decisions 51 or inferences about abstract variables 52 and sensory stimuli 53,54. Here, we were not interested in outcome uncertainty but rather in choice uncertainty, i.e. the uncertainty associated with using a particular decision strategy. Potentially, outcome and choice uncertainty share common neural substrates although this is not necessarily the case for conceptually different types of uncertainty 33.

Although we focused particularly on the hippocampus and the amygdala given their prominent role in animal studies on approach-avoidance conflicts, we note that in our task we observed similar effects for the DLPFC (particularly in the right hemisphere). DLPFC involvement is consistent with the well-described fMRI activity in this region for processing risks during choice 32.

The optimal policy also guided participants’ decisions as shown by behavioural model comparisons that pitted the optimal policy against a number of rather close-by competitors such as metrics related to immediate changes in expected value and policies that would have been optimal in slightly different tasks. We conjecture that humans may often fail to compute the optimal policy in more challenging real-life situations that entail larger action

repertoires, stricter time pressure, and the need to learn environmental states and contingencies 1. Previous research has not been able to address whether humans are capable of using elaborate decision policy for surmounting approach-avoidance conflicts. We could do so because our sequential decision-making task was formulated in the precise mathematical framework of a Markov decision process (which specifies criteria for optimality without committing to particular utility functions or risk preferences).

The optimal policy was related to a part of the DMPFC and the anterior cingulate cortex as well as IFG. The DMPFC and the IFG additionally tracked the choice uncertainty of the optimal policy. In general, these prefrontal regions are implicated in various types of decision-making tasks 31–36. The implication of the DMPFC and the anterior cingulate cortex might also point to further converging neural processes across species since medial prefrontal regions are also implicated in rodent approach-avoidance tasks 15. A close-up comparison indicates commonalities and differences in activation patterns with our previous study that used a foraging task without predation—and thus without approach-avoidance conflict 2 (Supplementary Figure 10). In both studies, choice according to the optimal policy was related to activity clusters in the anterior cingulate cortex (extending into the midcingulate cortex) which—although not overlapping—were located in close proximity. The choice uncertainty of the optimal policy was related to slightly more anterior DMPFC activity in the previous compared with the current study. In our view, the biggest difference between the studies concerns the role of the DMPFC extending into the dorsal part of the anterior cingulate cortex: In the previous study, activity in this regions robustly scaled with the discrepancy between using the heuristic—in that case the probability of foraging gain—and using the optimal policy. In the current study, the same region scaled with the optimal policy itself (and its choice uncertainty). This informal comparison between our studies may suggest the exciting possibility that the approach-avoidance conflict inherent in the choice options of the current task shifted the balance between the metrics encoded in the DMPFC—a conjecture to be tested in future experiments.

From a process perspective, we deem it intriguing that overlapping regions tracked the decision variable itself as well as its choice uncertainty. This seemed to be the case in the hippocampus, amygdala, DLPFC, DMPFC, and IFG. Choice uncertainties were calculated as the derivatives of the logistic functions linking the decision variables to participants' choices, which implies that choice uncertainties are highest at the inflection points of the logistic functions and lowest at the outer ranges when decision variables clearly prescribe one of the two actions. We speculate that overlapping representations of a decision variable and their choice uncertainty might aid in adjusting the degree to which these decision variables are used for choice, which hints potentially at local and distributed neural controllers for arbitrating between the heuristic and optimal policies (similar to what has been shown for model-free versus model-based learning 55,56).

Using a approach-avoidance conflict task with a precise decision-theoretic framework, we demonstrate that human DMPFC computes a sophisticated optimal policy on top of a predator probability heuristic that relates to the hippocampus and the amygdala—along with the associated choice uncertainty. These findings argue against a monolithic view of approach-avoidance conflicts and provide evidence for an interplay of two algorithms

implemented by multiple controllers. Our study dovetails both on a conceptual and on a neural level with work on survival circuits and risk assessment in rodents and humans 9,10,28,29,50. In particular, we link mathematically defined decision policies of varying complexity to flexible, higher-order cortical regions. Thereby, our study opens new avenues for translational research on the role of approach-avoidance conflicts for anxiety.

Methods

Participants

fMRI sample in Zurich—We recruited 29 participants via mailing lists of local universities. Five participants were excluded: One due to head motion > 4 mm during MRI, one due to an incidental medical finding revealed by MRI, and three who behaved almost deterministically, i.e., they selected one of the two choice options in more than 0.85 of the retained trials. (We selected 0.85 as cut-off during the analysis process. The proportion of choosing the foraging in remaining sample was 0.59 ± 0.10 , mean \pm SD; see Analyses and models of choice and reaction time data). The final sample comprised 24 participants (11 female; age = 25.0 ± 3.9 years). Due to equipment malfunction, two participants only performed nine out of ten sessions. Participants received a show-up fee of CHF 50 plus a variable amount (see Instructions and task).

Behavioural sample in Hamburg—We recruited 26 participants via a local online platform. Three participants were excluded: One due to equipment malfunction and two who selected one of the two choice options in more than 0.85 of the retained trials (value in remaining sample: 0.55 ± 0.18 , mean \pm SD, for the foraging option). The final sample comprised 23 participants (15 female; age = 25.9 ± 3.5 years). Due to time constraints, one participant only performed eight out of ten sessions. Participants received a show-up fee of EU 12 plus a variable amount.

The study was conducted in accord with the Declaration of Helsinki and approved by the governmental research ethics committees (fMRI sample: Kantonale Ethikkommission Zürich, KEK-ZH-Nr. 2013-0328; behavioural sample: Ethikkommission der Ärztekammer Hamburg, PV5746). All participants gave written informed consent using a form approved by the ethics committees.

Sequential decision-making task

Instructions and task—See Figure 1 for an overview of the task setup. Participants received detailed written and oral step-by-step instructions (see Supplementary Note 1: Task instructions), which presented the task as virtual foraging under the dual threats of predation and of starvation. To familiarize themselves with the task, participants performed two training sessions: A first short training session of four forests with five days each (after which participants could ask questions) and a second longer training session of 24 forests with five days each. The behavioural sample additionally received a questionnaire testing for task understating after the first training session (see Supplementary Note 2: Comprehension questionnaire, Supplementary Table 10). In our task, “forest” refers to a mini-block of “days” with each “day” being one trial. For the main behavioural task during fMRI

scanning, forests always lasted five days, but to maximize fMRI efficiency these five days could be interrupted, i.e. the first few days were played in the scanner, and the remaining days afterwards. Participants knew about this feature but did not know at which point a given forest would be interrupted. Participants performed ten sessions of the main behavioural task in the MR scanner (fMRI sample) or on a desktop computer (behavioural sample). The number of days per forest played in the scanner followed an exponential distribution with a mean of 2.5 resulting in 40 days in 24 forests per session. After fMRI scanning, participants completed one randomly selected forest per session, and were rewarded according to performance in these ten forests. For each forest in which they survived (regardless of the final energy state) participants received one additional reward point. On top of that, they received one reward point for each time within a forest that their energy level reached five points. Each point corresponded to CHF 1.50 (fMRI sample) or EUR 1 (behavioural sample). The task was presented using the MATLAB toolbox Cogent (www.vislab.ucl.ac.uk). After the experiment, all participants received a questionnaire asking for specific strategies and for ratings of task components (see Supplementary Note 3: Post-experiment questionnaire, Supplementary Table 10). For exploratory analyses, participants in the behavioural sample additionally filled in short forms of the State-Trait Anxiety Inventory (STAI) 57 and of the Need for Cognition (NFC) scale 58.

Mathematical framework and optimal policy—We modelled the task as a Markov decision process (MDP) 59. MDPs are specified by (1) the possible states, (2) the action repertoire, (3) the transition matrix between these states, (4) the rewards associated with transitions, and (5) the temporal horizon. In the following, we list these components:

- (1) States: 12 states per forest, i.e., 6 energy states (0-5 energy points) x 2 (weather types).
- (2) Actions: foraging or waiting.
- (3) Transition matrix: this is constructed from the probabilities of predator attack, the probabilities of foraging gain, gain and loss magnitudes, and the transition probability between the two weather types (which is always 0.5 and independent of the chosen action). Zero energy states are absorbing. Waiting leads to a sure loss of one point. Foraging can lead to an attack of the predator, which results in a transition to a zero energy state. If the predator does not attack, transitions depend on whether a foraging gain occurs or not. That is, the energy state is either increased by the gain magnitude (with total energy being capped at five) or reduced by two points.
- (4) Rewards: all transitions to zero energy states are associated with a reward of -1, all transitions to five energy points are associated with a reward of +1, and all other transitions with a reward of 0. That is, the rewards in the MDP reflect the monetary incentives participants had when performing the task.
- (5) Temporal horizon: our task imposes a finite time horizon of five steps from the start of each forest. Time steps are called days.

The optimal policy specifies actions that maximize obtained rewards. That is, the optimal policy depends on the choice at the current time step and the remaining—up to four—time steps. To derive optimal policies in our finite-horizon scenario, we used backward induction: Specifically, we started from the final time step (i.e., day five) and calculated the values of the two choice options (i.e., foraging or waiting) for each state. These values depend on the possible transitions from the respective states. If the value of foraging is higher than the value of waiting, foraging is the deterministically better option in that state and at that time step—or vice versa. If both choice options have the same value, the optimal choice is indifferent between the two options and this value is used to calculate the values for the second-to-last time step. If the two choice options differ in value, the value of the better choice option (i.e., the maximum over the values for foraging and waiting) is used to determine the optimal choice and to calculate the values for the second-to-last time step and. This procedure was repeated until arriving at the first time step (i.e., day one).

The optimal policy *per se* specifies the action to choose and thus does not allow for variability in the decision process (i.e., in some cases waiting and foraging entail large value differences whereas in other cases the two choice options have quite similar values). We therefore used the continuous value differences between the two choice options as predictors of participants' choices, RTs, and fMRI data. For brevity, we often use the term “optimal policy” to refer to the value differences between the foraging and waiting under the optimal policy.

The optimal policy applies to the scenario as instructed and incentivized in our experiment. It is a possibility that participants may actually have tried to solve different scenarios. We therefore tested four policies that are optimal under different scenarios. We refer to these policies as “pseudo-optimal.” One pseudo-optimal policy just considers a finite time horizon of one time step. A second pseudo-optimal policy neglects the transitions according to the predator probability (i.e., this probability is set to zero for calculating the corresponding pseudo-optimal policy). A third pseudo-optimal policy neglects the transitions according to the probability of foraging gain. For a fourth pseudo-optimal policy, which was added during the revision process, optimal policies were averaged for one to five days according to the number of times these horizons actually occurred in the main experiment (i.e., an average of 2.5 days, which was implemented to enhance fMRI design efficiency).

All calculations were carried out in MATLAB.

Choice uncertainties and discrepancy—We conjectured that choice uncertainties of the employed DVs and possibly the discrepancy between the DVs may be reflected in RT and fMRI data. We quantified uncertainties on the basis of the derivative of the mean across participants of the fitted logistic functions (Figure 2c, 2d). These derivatives capture the intuition that a small deviation in the DV (e.g., due to perceptual or computational error) has a large impact on the resulting choice at the inflection point of the logistic function (i.e., at the point at which participants are indifferent between the two choice options) than at the ranges of the decision variable where the values of the logistic function are close to zero (prescribing waiting) or close to one (prescribing foraging). Additionally, to quantify the

discrepancy in the prescriptions of the two employed DVs, we took the absolute differences in the mean of the two fitted logistic functions across participants.

Analyses and models of choice and reaction time data

Participants received a total of 400 trials, i.e., days (except for two participants who only received 360 trials due to equipment malfunction). On average, participants were “alive”—and were thus able to make decisions—in 354.4 ± 16.1 trials, mean \pm SD, (fMRI sample) and 354.4 ± 21.9 trials (behavioural sample). Of these, they failed to answer in 8.4 ± 8.8 trials (fMRI sample) and 5.4 ± 8.3 trials (behavioural sample), which left 346.0 ± 17.6 trials (fMRI sample) and 349.0 ± 27.6 trials (behavioural sample) for analyses.

To explain participants’ choices, i.e., their probability of choosing the foraging option, P_{forage} , we used logistic regression models (implemented in the MATLAB function *mnrfit*) of the following generic form:

$$P_{\text{forage}} = 1/(1 + \exp(-DV))$$

with the following form of the decision variable (DV):

$$DV = \beta_0 + \beta_1 * predictor$$

As *predictor*, we first considered the 16 variables listed in Table 1. In the same vein, we tested models including two predictors:

$$DV = \beta_0 + \beta_1 * predictor_1 + \beta_2 * predictor_2$$

We also tested interaction models:

$$DV = \beta_0 + \beta_1 * predictor_1 + \beta_2 * predictor_2 + \beta_3 * predictor_1 * predictor_2$$

For each model we approximated model evidence by calculating the Bayesian Information Criterion (BIC), which penalizes model complexity:

$$BIC = SDR + k * \ln(n)$$

where SDR is the sum of deviance residuals, i.e., twice the difference between the maximum achievable log likelihood and that attained under the fitted model as given by *mnrfit*, k is the total number of predictors including the intercept, and n is the number of data points per participant.

We performed both fixed-effects and random-effects analyses. The latter assume that different participants may use different models. We used the Bayesian model selection procedure implemented in SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/>) to calculate protected exceedance probabilities, which measure the likelihood that any given model is more

frequent than all other models in the comparison set 60. For the Bayesian model selection procedure, BIC values were multiplied by -0.5 to scale values with respect to the appropriate conventions 61,62.

We analysed log-transformed RTs using linear mixed effects models as implemented in the R package lmer (<http://cran.r-project.org/web/packages/lme4/index.html>) 63. Random effects for participants included a random intercept and random slopes for all variables. Since models of choice data did not provide evidence for the suitability of interaction terms, we did not include any interaction terms as fixed- or random-effects in the models of RT data. P-values and degrees of freedom were derived using the R package lmerTest (<https://cran.r-project.org/web/packages/lmerTest/>). Log-likelihood differences were calculated between the models including all fixed effects relative to the models without the respective fixed effect (but with the same random-effects structure).

FMRI data acquisition

Data were acquired on a 3 T (Philips Achieva, Best, The Netherlands) MR scanner using a 32-channel head coil. Functional images were recorded using a T2*-weighted echo-planar imaging (EPI) sequence (TR 2.1 s; TE 30 ms; flip angle 80°). A total of 37 axial slices were sampled for whole brain coverage (matrix size 96 × 96; in-plane resolution 2.5 × 2.5 mm²; slice thickness 2.8 mm; 0.5 mm gap between slices; slice tilt 0°). Functional images were collected in ten sessions of 170 volumes each. To obtain steady-state longitudinal magnetization, the first five volumes of each session were discarded. Field maps were acquired with a double echo gradient echo field map sequence, using 32 slices covering the whole head (TR 349.11 ms; TE 4.099 and 7.099 ms; matrix size, 80 × 80; in-plane resolution 3 × 3 mm²; slice thickness 3 mm; 0.5 mm gap between slices; slice tilt 0°). Anatomical images were acquired using a T1-weighted scan (FoV 255 × 255 × 180 mm; voxel size 1 × 1 × 1 mm³).

FMRI data analyses

All fMRI analyses were performed in SPM12. The FieldMap toolbox was used to correct for geometric distortions caused by susceptibility-induced field inhomogeneities 64. Preprocessing of EPI data included rigid-body realignment to correct for head movement, unwarping, and slice time correction. EPI images were then coregistered to the individual's T1 weighted image using a 12-parameter affine transformation and normalized to the Montreal Neurological Institute (MNI) T1 reference brain template using the extended unified segmentation algorithm in SPM12 65. Normalized images were smoothed with an isotropic 8 mm full width at half-maximum Gaussian kernel. The six motion correction parameters estimated from the realignment procedure were entered as covariates of no interest. Regressors were convolved with the canonical HRF and low frequency drifts were excluded using a high-pass filter with a 128 s cut-off.

In the general linear model (GLM) the three distinct phases of the task (forest, choice, and outcome phases; see Figure 1) were entered as events with a duration of 0 s (i.e., as stick functions). Choice and outcome phases in which participants had starved or for which they did not reply were not explicitly modelled. We were mostly interested in the choice phase

and ran a primary GLM with a combination of variables that emerged in our analyses of behavioural and RT data (see Figure 3 and Table 3). That is, our primary GLM included the six parametric modulators of the choice phase: (1) the predator probability, (2) the DV under the optimal policy, (3&4) the associated choice uncertainties of the former two metrics, (5) the discrepancy between these two metrics, and (6) log-transformed RTs. We report analyses in which parametric modulators competed for variance (i.e., without serial orthogonalization). We also ran five separate GLMs with serial orthogonalization such that each of the five parametric modulators was entered last in one GLM. These GLMs revealed the same clusters and same values as our primary GLM except for some numerical differences due to rounding. Due to collinearities between parametric modulators, there may be some BOLD activity that could be explained by several parametric modulators, which is not reported here, since we were interested in the specific effects of the respective parametric modulators. The forest phase was parametrically modulated by the current energy state. The outcome phase was parametrically modulated by the change in energy state (i.e., energy state at outcome minus energy state at choice).

In a secondary GLM, participants' choices were included as an additional parametric modulator (i.e., as a binary modulator coding for waiting or foraging; see Supplementary Table 13). In a tertiary GLM, choice uncertainties and the discrepancy were calculated from the independent behavioural sample (see Supplementary Table 14). For follow-up region of interest (ROI) analyses, we set up a GLM with four separate onset regressors for the four levels of predator probability (Figure 4) and three GLMs on the basis of the primary GLM, in which we additionally entered one of the three next-best DVs from the behavioural model comparison as parametric modulators (i.e., the probability of foraging gain, continuous energy, and expected energy change; see Supplementary Table 16 for the respective results). In two further GLMs, we added either the interaction of predator probability with the choice uncertainty of predator probability or the interaction of the optimal policy with the choice uncertainty of the optimal policy as parametric modulators (Supplementary Table 17).

We performed a second-level one-sample t-test (one-sided) on contrast images from all participants. All reported clusters are family-wise error (FWE) corrected for multiple comparisons at $p < 0.05$ using the SPM random field theory based approach. The cluster-defining threshold was $p < 0.001$. At this voxel-inclusion threshold the random-field theory approach in SPM correctly controls the false positive rate 66. ROI analyses were conducted using the toolboxes rfxplot and marsbar.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Giuseppe Castegnetti, Saurabh Khemka, Matthias Staib, Athina Tzovara, and Cristian Ioan for discussions and help with data acquisition. The Wellcome Trust Centre for Neuroimaging is supported by a strategic grant from the Wellcome Trust [091593/Z/10/Z]. CWK was supported by two grants from the German Research Foundation (DFG) during the final stages of manuscript preparation: the collaborative research centre SFB TRR 169 and an Emmy Noether Research Group [392443797]. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Bach DR, Dayan P. Algorithms for survival: a comparative perspective on emotions. *Nat Rev Neurosci.* 2017; 18:311–319. [PubMed: 28360419]
2. Korn CW, Bach DR. Heuristic and optimal policy computations in the human brain during sequential decision-making. *Nat Commun.* 2018; 9
3. Korn CW, Bach DR. Maintaining Homeostasis by Decision-Making. *PLoS Comput Biol.* 2015; 11
4. Huys QJM, et al. Interplay of approximate planning strategies. *Proc Natl Acad Sci U S A.* 2015; 112:3098–103. [PubMed: 25675480]
5. Huys QJM, et al. Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput Biol.* 2012; 8
6. Keramati M, Smittenaar P, Dolan RJ, Dayan P. Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proc Natl Acad Sci.* 2016; 113:12868–12873. [PubMed: 27791110]
7. Keramati M, Dezfouli A, Piray P. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLOS Comput Biol Biol.* 2011; 7:e1002055.
8. LeDoux J, Daw ND. Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nat Rev Neurosci.* 2018; 19:269–282. [PubMed: 29593300]
9. Mobbs D, Trimmer PC, Blumstein DT, Dayan P. Foraging for foundations in decision neuroscience: insights from ethology. *Nat Rev Neurosci.* 2018:1–9.
10. Mobbs D. The ethological deconstruction of fear(s). *Curr Opin Behav Sci.* 2018; 24:32–37.
11. Griebel G, Holmes A. 50 years of hurdles and hope in anxiolytic drug discovery. *Nat Rev Drug Discov.* 2013; 12:667–87. [PubMed: 23989795]
12. Haller J, Aliczki M, Gyimesine Pelczer K. Classical and novel approaches to the preclinical testing of anxiolytics: A critical evaluation. *Neurosci Biobehav Rev.* 2013; 37:2318–2330. [PubMed: 22981935]
13. Cryan JF, Sweeney FF. The age of anxiety: Role of animal models of anxiolytic action in drug discovery. *Br J Pharmacol.* 2011; 164:1129–1161. [PubMed: 21545412]
14. Kirlic N, Young J, Aupperle RL. Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behav Res Ther.* 2017; 96:14–29. [PubMed: 28495358]
15. Calhoun GG, Tye KM. Resolving the neural circuits of anxiety. *Nat Neurosci.* 2015; 18:1394–1404. [PubMed: 26404714]
16. Gray, JA, McNaughton, N. *The neuropsychology of anxiety: An enquiry into the functions of the septohippocampal system.* Oxford University Press; 2000.
17. Aupperle RL, Melrose AJ, Francisco A, Paulus MP, Stein MB. Neural substrates of approach-avoidance conflict decision-making. *Hum Brain Mapp.* 2015; 36:449–462. [PubMed: 25224633]
18. Bach DR, et al. Human hippocampus arbitrates approach-avoidance conflict. *Curr Biol.* 2014; 24:541–7. [PubMed: 24560572]
19. Loh E, et al. Parsing the role of the hippocampus in approach–avoidance conflict. *Cereb Cortex.* 2016; :1–15. DOI: 10.1093/cercor/bhw378 [PubMed: 25139941]
20. O’Neil EB, et al. Examining the role of the human hippocampus in approach-avoidance decision making using a novel conflict paradigm and multivariate functional magnetic resonance imaging. *J Neurosci.* 2015; 35:15039–15049. [PubMed: 26558775]
21. Ito R, Lee ACH. The role of the hippocampus in approach-avoidance conflict decision-making: Evidence from rodent and human studies. *Behav Brain Res.* 2016; 313:345–357. [PubMed: 27457133]
22. Schlund MW, et al. The tipping point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *Neuroimage.* 2016; 136:94–105. [PubMed: 27153979]
23. Mobbs D, et al. When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science (80-).* 2007; 317:1079–1083.
24. Qi S, et al. How cognitive and reactive fear circuits optimize escape decisions in humans. *Proc Natl Acad Sci.* 2018; 115:3186–3191. [PubMed: 29507207]

25. Bach DR. Anxiety-like behavioural inhibition is normative under environmental threat-reward correlations. *PLoS Comput Biol*. 2015; 11:1–20.
26. Korn CW, et al. Amygdala lesions reduce anxiety-like behavior in a human benzodiazepine-sensitive approach–avoidance conflict test. *Biol Psychiatry*. 2017; 82
27. Khemka S, Barnes G, Dolan RJ, Bach DR. Dissecting the function of hippocampal oscillations in a human anxiety model. *J Neurosci*. 2017; 37:6869–6876. [PubMed: 28626018]
28. McNaughton N, Corr PJ. Survival circuits and risk assessment. *Curr Opin Behav Sci*. 2018; 24:14–20.
29. Blanchard DC. Risk assessment: at the interface of cognition and emotion. *Curr Opin Behav Sci*. 2018; 24:69–74.
30. Amemori K, Graybiel AM. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat Neurosci*. 2012; 15:776–85. [PubMed: 22484571]
31. Symmonds M, Wright ND, Bach DR, Dolan RJ. Deconstructing risk: separable encoding of variance and skewness in the brain. *Neuroimage*. 2011; 58:1139–49. [PubMed: 21763444]
32. Mohr PNC, Biele G, Heekeren HR. Neural processing of risk. *J Neurosci*. 2010; 30:6613–6619. [PubMed: 20463224]
33. Bach DR, Dolan RJ. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat Rev Neurosci*. 2012; 13:572–86. [PubMed: 22781958]
34. Bartra O, McGuire JT, Kable JW. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*. 2013; 76:412–427. [PubMed: 23507394]
35. Clithero JA, Rangel A. Informatic parcellation of the network involved in the computation of subjective value. *Soc Cogn Affect Neurosci*. 2014; 9:1289–1302. [PubMed: 23887811]
36. Rushworth MFS, Noonan MP, Boorman ED, Walton ME, Behrens TE. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 2011; 70:1054–1069. [PubMed: 21689594]
37. Kolling N, Behrens TEJ, Mars RB, Rushworth MFS. Neural mechanisms of foraging. *Science*. 2012; 336:95–8. [PubMed: 22491854]
38. Kolling N, Wittmann M, Rushworth MFS. Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*. 2014; 81:1190–202. [PubMed: 24607236]
39. Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci*. 2011; 14:933–9. [PubMed: 21642973]
40. Mata R, Wilke A, Czienskowski U. Foraging across the life span: is there a reduction in exploration with aging? *Front Neurosci*. 2013; 7:53. [PubMed: 23616741]
41. Shenhav A, Straccia MA, Cohen JD, Botvinick MM. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat Neurosci*. 2014; 17:1249–1254. [PubMed: 25064851]
42. Constantino SM, Daw ND. Learning the opportunity cost of time in a patch-foraging task. *Cogn Affect Behav Neurosci Behav Neurosci*. 2015; 15:837–53.
43. Pearson JM, Watson KK, Platt ML. Decision making: The neuroethological turn. *Neuron*. 2014; 82:950–965. [PubMed: 24908481]
44. Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu Rev Psychol*. 2011; 62:451–82. [PubMed: 21126183]
45. Gu X, FitzGerald THB. Interoceptive inference: homeostasis and decision-making. *Trends Cogn Sci*. 2014; 18:269–70. [PubMed: 24582825]
46. Fawcett TW, et al. The evolution of decision rules in complex environments. *Trends Cogn Sci*. 2014; 18:153–61. [PubMed: 24467913]
47. Dayan P. Rationalizable irrationalities of choice. *Top Cogn Sci*. 2014; 6:204–228. [PubMed: 24648392]
48. Tovote P, Fadok JP, Lüthi A. Neuronal circuits for fear and anxiety. *Nat Rev Neurosci*. 2015; 16:317–331. [PubMed: 25991441]
49. Jimenez JC, et al. Anxiety cells in a hippocampal-hypothalamic circuit. *Neuron*. 2018; 97:670–683.e6. [PubMed: 29397273]

50. Blanchard DC. Translating dynamic defense patterns from rodents to people. *Neurosci Biobehav Rev.* 2017; 76:22–28. [PubMed: 28434585]
51. Payzan-LeNestour E, Dunne S, Bossaerts P, O’Doherty J. The neural representation of unexpected uncertainty during value-based decision making. *Neuron.* 2013; 79:191–201. [PubMed: 23849203]
52. Rigoli F, Michely J, Friston KJ, Dolan RJ. The role of the hippocampus in weighting expectations during inference under uncertainty. *Cortex.* 2019; 115:1–14. [PubMed: 30738997]
53. Harrison LM, Duggins A, Friston KJ. Encoding uncertainty in the hippocampus. *Neural Networks.* 2006; 19:535–546. [PubMed: 16527453]
54. Strange BA, Duggins A, Penny W, Dolan RJ, Friston KJ. Information theory, novelty and hippocampal responses: Unpredicted or unpredictable? *Neural Networks.* 2005; 18:225–230. [PubMed: 15896570]
55. Lee SW, Shimojo S, O’Doherty JP. Neural computations underlying arbitration between model-based and model-free learning. *Neuron.* 2014; 81:687–699. [PubMed: 24507199]
56. Boureau Y-L, Sokol-Hessner P, Daw ND. Deciding how to decide: Self-control and meta-decision making. *Trends Cogn Sci.* 2015
57. Laux, L, Glanzmann, P, Schaffner, P, Spielberger, CD. STAI - State-Trait-Angstinventar. Theoretische Grundlagen und Handanweisung [in German]. Beltz Test GmbH; 1981.
58. Bless H, Wänke M, Bohner G, Fellhauer RF, Schwarz N. Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben. *Zeitschrift für Sozialpsychologie.* 1994; :147–154. DOI: 10.1111/j.1467-839X.2008.00269.x
59. Sutton, RS, Barto, AG. Reinforcement learning: An introduction. MIT Press; 1998.
60. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian model selection for group studies - revisited. *Neuroimage.* 2014; 84:971–85. [PubMed: 24018303]
61. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *Neuroimage.* 2009; 46:1004–17. [PubMed: 19306932]
62. Penny WD. Comparing dynamic causal models using AIC, BIC and free energy. *Neuroimage.* 2012; 59:319–330. [PubMed: 21864690]
63. Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang.* 2008; 59:390–412.
64. Hutton C, et al. Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage.* 2002; 16:217–240. [PubMed: 11969330]
65. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage.* 2005; 26:839–851. [PubMed: 15955494]
66. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci.* 2016; 113:7900–7905. [PubMed: 27357684]

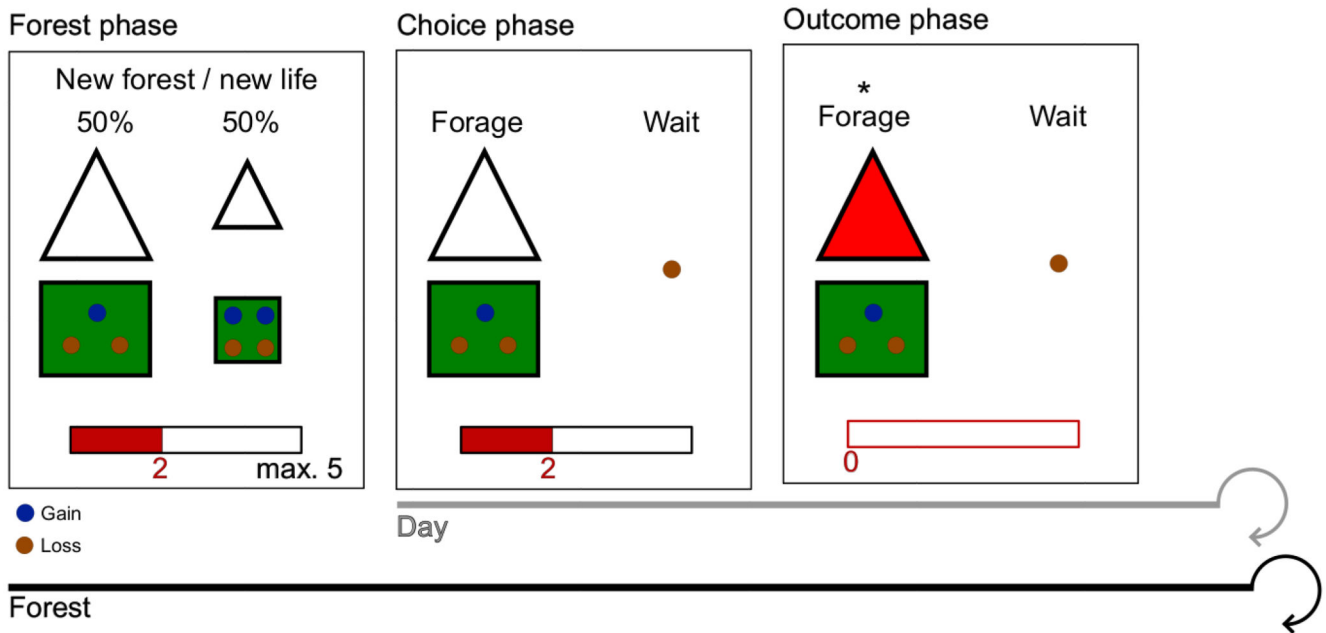


Figure 1. Task outline.

Participants performed a sequential decision-making task that was framed as foraging under the dual threats of predation and starvation. Current energy was depicted as an energy bar. Participants were monetarily incentivized to keep their energy above zero and to attain the maximum energy state of five as often as possible (intermediate energy states did not directly translate into monetary payoffs). For each mini-block of trials, called “forest,” energy was reset and foraging options varied. During the initial *forest phase* (3.5 sec), participants were informed about their initial energy (2, 3, or 4 energy points) and the two possible foraging options in the current forest over five trials, called “days.” The number of days was not depicted and varied according to an exponential distribution (for fMRI design efficiency) but participants knew that their final payoffs would depend on a subset of forests for which they would complete exactly five days. Triangles of four different sizes showed the current predator probability (0.1, 0.2, 0.3, or 0.4). Green rectangles of four different sizes depicted the probability of foraging gain (0.2, 0.4, 0.6, or 0.8). The number of blue dots within the green rectangles showed gain magnitudes (varying from 0 to 4). Dark red dots showed losses, which were set to 1 for the waiting option (with a probability of 1) and set to 2 for the foraging option (with a probability of $1 - (\text{probability of foraging gain})$). After a fixation interval (3.5 sec) the *choice phase* started (2.0 sec): One of the two foraging options was presented on the screen (with a probability of 0.5) and participants had to make a decision between waiting and this foraging option. Sides were counterbalanced. If participants failed to respond, the words “Too slow” appeared. In the example shown, the participant chose foraging (indicated by an asterisk). After an interval of 1.0 sec, the *outcome phase* started (1.0 sec). In the current example, the predator attacked and all energy points were lost. After a variable fixation interval (between 0.5 and 3.8 sec), a new day or a new forest was depicted.

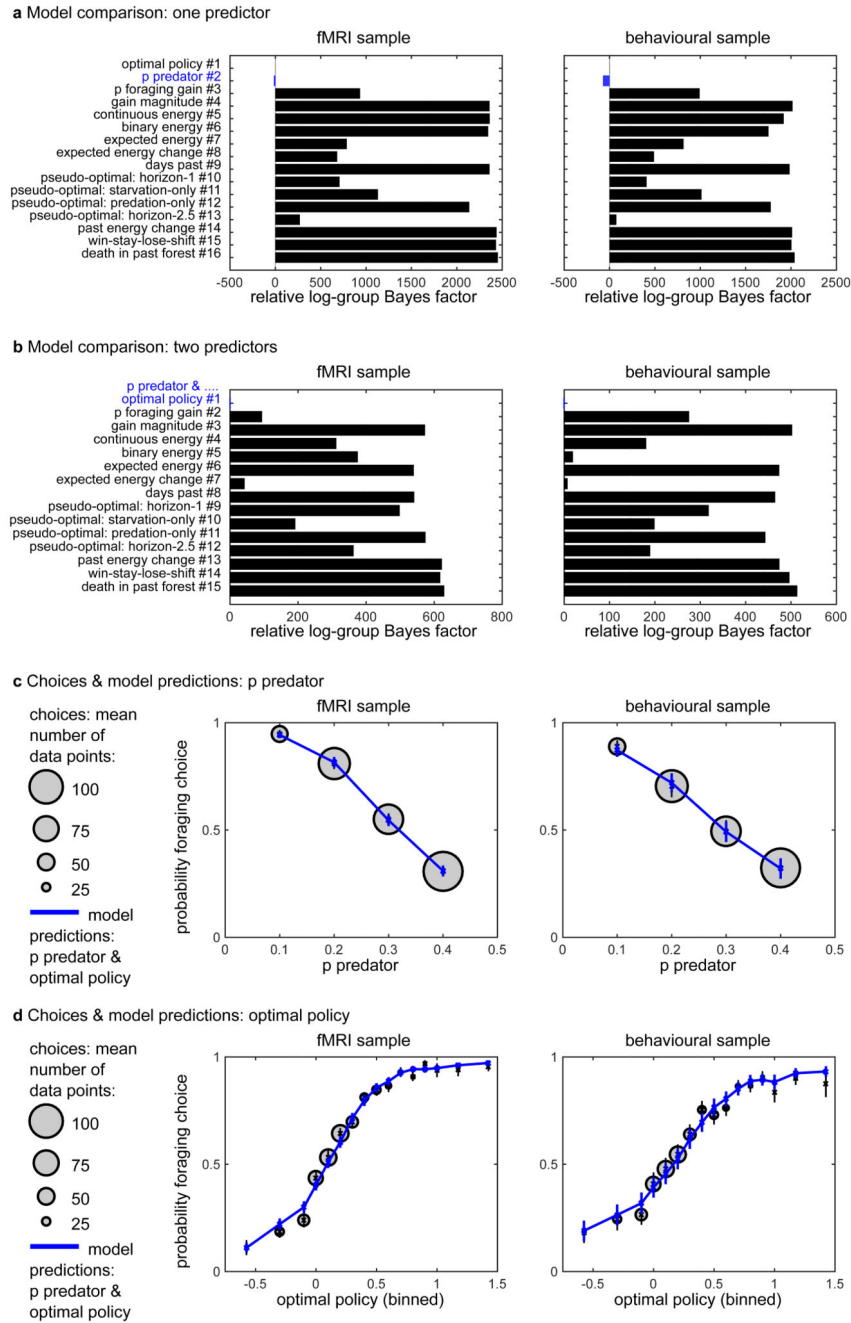


Figure 2. Models of choice data in the fMRI sample.

- (a) Bayesian Model comparisons show that the probability of predator was the best single predictor of participants' choices. The plot depicts fixed-effects analyses using relative log-group Bayes factors based on Bayesian Information Criterion (BIC) relative to model #1.
- (b) A model that additionally included the optimal policy best explained remaining variance in participants' choices.
- (c) The winning model captures the relationship between participants' average choices and the predator probability.

(d) The winning model captures the relationship between participants' average choices and the optimal policy (binned value differences of foraging versus waiting). Number of participants in the fMRI sample, $n=24$. Number of participants in the behavioural sample, $n=23$. Better fit is indicated by smaller log-group Bayes factors (i.e., larger negative values). In the right-hand panels error bars are SEM. Per data bin, circles depict mean empirical data points and lines and crosses depict mean model predictions (averaged for simulated data according to each participant's model fit). See Table 1 for a list that specifies the considered decision variables (DVs) for the task and thus the models tested here. See Supplementary Tables 1-3 for detailed model comparisons in the fMRI sample and Supplementary Tables 4-6 for detailed model comparisons in the behavioural sample. Supplementary Tables 7-9 present shared variances between the DVs and confusion matrices. See Supplementary Figure 1 for the relationships among the 16 DVs included in the models. See Supplementary Figures 2, 3 for plots showing that the winning model captures the data split according to the other 14 DVs and that the winning model makes better qualitative predictions than the other models considered. Supplementary Figure 4 depicts individual variability in the fMRI sample.

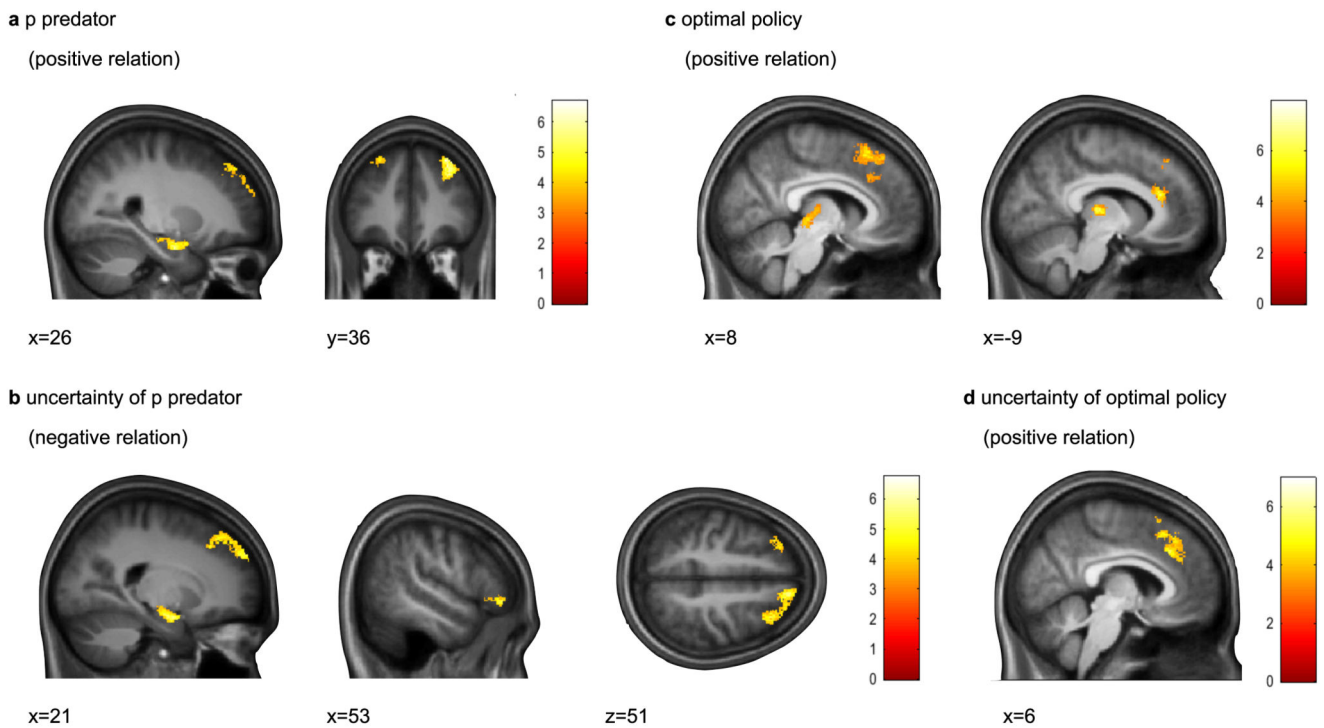


Figure 3. FMRI results during the choice phase.

(a) Predator probability showed a positive relation within a cluster spanning right anterior hippocampus and amygdala as well as within bilateral DLPFC.

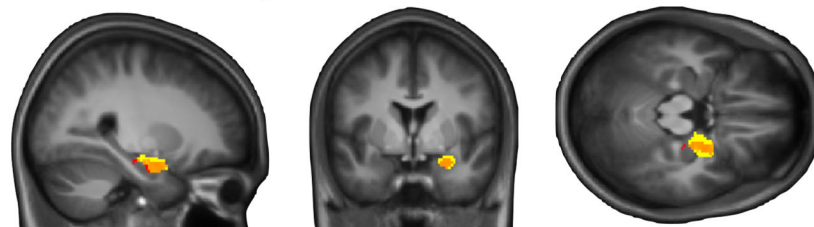
(b) The choice uncertainty according to the predator probability showed a negative relation in right anterior hippocampus extending into amygdala as well as bilateral DLPFC and right lateral IFG.

(c) The optimal policy showed a positive relation within DMPFC, extending into pre-SMA and ACC, and thalamus among other regions.

(d) The choice uncertainty according to the optimal policy showed a positive relation in DMPFC extending into pre-SMA.

Number of participants, $n=24$. Overlay on group average T1-weighted image in MNI space; clusters are whole-brain family-wise error (FWE) corrected for multiple comparisons at $p < 0.05$ with a cluster-defining threshold of $p < 0.001$. See Table 3 for a list of all clusters. See Supplementary Table 11 for the relationships among the variables included as parametric modulators during the choice phase. See Figure 4 for the overlap of the hippocampus/amygdala clusters depicted in **(a)** and **(b)**. Supplementary Figure 5 visualizes the overlap of all clusters. See Supplementary Table 12 and Supplementary Figure 6 for fMRI results during the outcome phase. See Supplementary Table 13 for the results of a secondary model that additionally included participants' choices as a parametric modulator. See Supplementary Table 14 for the results of a tertiary model in which choice uncertainties were derived from the behavioural sample. Supplementary Table 15 and Supplementary Figure 8 present the results of a covariate analyses testing for inter-individual differences. Supplementary Table 16 and Supplementary Figures 8, 9 present non-independent region of interest analyses. See Supplementary Table 17 for contrasts testing for interaction effects.

Supplementary Figure 10 visualizes the overlap between clusters from the current study and clusters from a related previous study 2.

a Overlap of clusters in hippocampus/amygdala

x=26

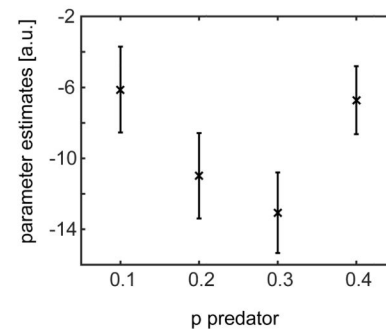
y=0

z=-19

p predator (positive relation)

uncertainty of p predator (negative relation)

overlap

b Parameter estimates in overlap ROI (depicted in a)**Figure 4. Visualization of the clusters in the hippocampus extending into the amygdala.**

(a) Overlap of the functional clusters for predator probability *per se* (red) and for the choice uncertainty of predator probability (yellow; see also depiction of these clusters in Figure 3a, 3b). Comparing these functional clusters to anatomical masks of the Automated Anatomical Labeling (AAL) atlas showed that of the 340 voxels in the cluster for predator probability *per se* 194 were in the hippocampus and 113 in the amygdala. Of the 672 voxels in the cluster for uncertainty of predator probability 294 were in the hippocampus and 270 in the amygdala.

(b) Visualization of the relation between the parameter estimates in the overlap region of interest (ROI) depicted in **(a)**. Parameter estimates were derived from a follow-up GLM, in which the four levels of predator probability were modelled as separate onset regressors for the choice phase. Notably, the predator probabilities used here cover a different range than that used in a previous study¹⁸ which showed a purely linear impact of predator probability on threat probability (with threat probabilities: 0.2, 0.5, 0.8). Please note that the number of data points was not equally distributed for the four bins. Parameter estimates were extracted using the toolbox marsbar.

Number of participants, $n=24$.

Table 1
List of variables used in behavioural model comparisons, RT analyses, and fMRI analyses

Variable	Description
#1 optimal policy	The optimal decision variable (DV) that takes into account all remaining time points (days), energy states, and the transition probabilities between them. Therefore, optimal policy cannot be directly inferred from information on the screen but necessitates rather complex computations.
#2 predator probability	The probability of the predator attacking when the foraging option is chosen. This probability is depicted by the size of a triangle on the screen and takes the values of 0.1, 0.2, 0.3, or 0.4. A predator attack leads to immediate death (i.e., an energy state of zero).
#3 probability of foraging gain	The probability of obtaining the depicted number of gains when the foraging option is chosen. This probability is depicted by the size of a green rectangle on the screen and takes the values of 0.2, 0.4, 0.6, or 0.8.
#4 gain magnitude	The number of points added to the energy state if foraging is successful and the predator does not attack. Gains vary from zero to four and are depicted by the number of blue dots within the green rectangle. Waiting entails a sure loss of one energy point and unsuccessful foraging always entails a constant loss magnitude of two points (depicted as red dots).
#5 continuous energy	Energy varies on a continuous scale from one to five in steps of one point. Zero energy corresponds to being dead (and thus no choice can be made in a zero energy state).
#6 binary energy	An energy state of one is special because in that state waiting leads to sure death. The “binary energy” variable therefore distinguishes between an energy state of one and all other energy states.
#7 expected energy	The foraging option entails an expected energy state, in the sense of expected value. This metric is thus calculated from the predator probability, the probability of foraging gain, the obtainable gain magnitude, and the continuous energy state (as well as the constant loss magnitude). The expected energy variable only takes one time step into account.
#8 expected energy change	The difference between the current energy state and the expected energy state for foraging (see #7).
#9 days past	The number of days (i.e., time steps) already spent in a given forest. The maximum number of days within a forest is always five.
#10 pseudo-optimal: horizon-1	This policy is optimal in the final time step (i.e., when only one day is left within a forest). Otherwise it can be regarded as pseudo-optimal because it is too short-sighted.
#11 pseudo-optimal: starvation-only	This policy would be optimal if no predators were present.
#12 pseudo-optimal: predation-only	This policy would be optimal if starvation were not possible.
#13 pseudo-optimal: horizon-2.5	This policy would be optimal if participants would not have been rewarded according to a full horizon of 5 days but according to an average horizon of 2.5 days, which was implemented in the main experiment to enhance fMRI design efficiency.
#14 past energy change	The difference between the energy states between choice and outcome phases of the past trial. Due to the Markov property of the task this past change is irrelevant for the optimal policy. This same metric can be evaluated during the outcome phase of each trial and signals how many energy points are gained or lost in the given trial. This variable is thus included as a parametric modulator during the outcome phase.
#15 win-stay-lose-shift	This DV prescribes foraging if the energy state increased with respect to the past trial and waiting if the energy state decreased. Win-stay-lose-shift is a binarized version of “past energy change.”
#16 death in past forest	Binary variable indicating whether participants reached zero energy points in the forest immediately prior to the current forest.
#17 uncertainty of p predator	When the prescriptions of the employed heuristic policy, i.e., of p predator, are closer to 0 (i.e., waiting) or 1 (i.e., foraging) uncertainty is lower than when the prescriptions lie in-between. Uncertainty is indexed by the derivative of the mean of the logistic function for p predator (see Figure 2c).
#18 uncertainty of optimal policy	When the absolute value differences according to the optimal policy are large (i.e., either clearly prescribing waiting or foraging) uncertainty is lower than when the absolute value differences are small (i.e., the optimal policy is more or less indifferent). Uncertainty is indexed by the derivative of the mean of the logistic function for the optimal policy (see Figure 2d).
#19 discrepancy in choice probabilities between p predator & optimal policy	In some cases, the heuristic policy of using p predator and the optimal policy make quite distinct prescriptions (high discrepancy), whereas in others they make quite similar prescriptions (low discrepancy). Discrepancy is indexed by the absolute differences of two logistic functions (see Figure 2c versus Figure 2d).

Variables #1 to #16 were used in models of choice behaviour. On the basis of the two variables in the behavioural model (i.e., on the basis of #1 and #2), we derived variables #17 to #19 and included these in RT and fMRI analyses.

Table 2
Reaction time data: Linear mixed effects model

Log-transformed RTs were analysed using linear mixed effects models; as implemented in the R package lmer. P-values and degrees of freedom were derived using the R package lmerTest. Significant effects are printed in bold font. Log-likelihood differences were calculated between the models including all fixed effects relative to the models without the respective fixed effect (but with the same random-effects structure). Better fit is indicated by smaller log-likelihood differences (i.e., larger negative values).

Predictor	estimates	degrees of freedom	t-values	p-values	log-likelihood difference	95%-confidence interval	
						lower limit	upper limit
fMRI sample (n=24)							
Intercept	6.66	24.45	147.49	< 0.001		6.56	6.75
#1 p predator	0.02	27.07	0.27	0.786	-0.1	-0.15	0.20
#2 optimal policy	-0.06	30.67	-3.04	0.005	-4.0	-0.10	-0.02
#3 uncertainty of p predator	0.55	25.17	3.43	0.002	-4.8	0.22	0.87
#4 uncertainty of optimal policy	0.12	43.17	1.73	0.091	-1.4	-0.02	0.27
#5 discrepancy between p predator & optimal policy	0.20	28.85	5.53	< 0.001	-10.1	0.13	0.27
Behavioural sample (n=23)							
Intercept	6.69	23.05	130.17	< 0.001		6.58	6.80
#1 p predator	-0.00	25.00	-0.04	0.970	0.0	-0.17	0.16
#2 optimal policy	-0.05	35.64	-3.03	0.005	-4.1	-0.09	-0.02
#3 uncertainty of p predator	0.31	28.31	2.69	0.012	-3.2	0.07	0.54
#4 uncertainty of optimal policy	-0.07	31.67	-0.85	0.400	-1.1	-0.22	0.09
#5 discrepancy between p predator & optimal policy	0.12	22.34	3.56	0.002	-4.9	0.05	0.20

Table 3
FMRI results during choice phase (primary GLM)

Clusters are whole-brain FWE corrected for multiple comparisons at $p < 0.05$ with a cluster-defining threshold of $p < 0.001$. Number of participants, $n=24$.

	Side	Peak voxel MNI coordinates (mm)			Cluster size (Voxel)	Peak t score
		X	Y	Z		
#1 p predator: positive (trial-by-trial relation with higher numbers indicating higher probability of predator attacking)						
Dorsolateral prefrontal cortex (DLPFC)	R	30	36	48	1325	6.69
Medial occipital cortex	L	-14	-74	-9	1112	6.17
Anterior hippocampus extending into amygdala	R	26	-6	-23	340	6.17
DLPFC	L	-29	33	50	194	4.75
#1 p predator: negative						
Inferior occipital gyrus	L	-29	-87	-12	462	6.48
Inferior occipital gyrus	R	21	-92	-9	294	6.05
#2 optimal policy: positive (trial-by-trial relation with higher numbers indicating value of the foraging option versus the waiting option according to the optimal policy)						
Thalamus	L	-8	-21	6	250	7.92
Inferior frontal gyrus (IFG) extending into insula	R	32	23	5	950	7.58
Anterior cingulate cortex	L	-9	32	20	260	6.59
Medial occipital cortex	L	-23	-72	-9	417	6.25
Thalamus	R	11	-27	-6	408	5.69
Posterior dorsal medial prefrontal cortex (DMPFC) extending into supplementary motor area	L/R	8	23	53	1393	5.57
IFG extending into insula	L	-30	21	-8	152	4.56
#2 optimal policy: negative						
Superior parietal gyrus	L	-17	-57	69	472	7.41
#3 uncertainty of p predator: positive (trial-by-trial relation with higher numbers indicating higher choice uncertainty according to the predator probability)						
None						
#3 uncertainty of p predator: negative						
DLPFC extending into DMPFC	R	14	38	54	1977	6.74
Posterior middle temporal gyrus	R	59	-56	-5	197	6.10
Lateral IFG	R	53	39	-6	572	6.06
Anterior hippocampus extending into amygdala	R	21	-9	-18	672	5.94
Medial occipital cortex	L	-14	-87	32	145	5.00
DLPFC	L	-32	29	51	149	4.83
#4 uncertainty of optimal policy: positive (trial-by-trial relation with higher numbers indicating higher choice uncertainty according to the optimal policy)						
DLPFC & DMPFC	R	17	12	57	2484	6.96
Insula	R	21	23	-5	707	6.37
DLPFC	R	39	35	41	935	6.16
DLPFC extending lateral IFG	R	41	54	12	800	5.36
#4 uncertainty of optimal policy: negative						
Posterior cingulate cortex	L	-8	-51	33	149	4.84

	Side	Peak voxel MNI coordinates (mm)			Cluster size (Voxel)	Peak t score
		X	Y	Z		
#5 discrepancy between p predator & optimal policy: positive (trial-by-trial relation with higher numbers indicating larger discrepancies)						
Thalamus	L	-9	-9	2	164	5.34
#5 discrepancy between p predator & optimal policy: negative						
None						