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

Christoph W. Korn and Dominik R. Bach

Jointly minimizing multiple threats over extended time horizons enhances survival. Consequently, many tests of approach–avoidance conflicts 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 the 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.

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. 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. However, 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. 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 that implement a variety of decision policies spanning different levels of sophistication and efficacy.

Several of these policies and neural controllers may be concurrently invoked when seeking a solution to a complex situation that threatens survival. A particularly relevant example for such decision problems is foraging for food under the risk of predation, which involves an approach–avoidance conflict (that is, 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. Some of these paradigms have been successfully reverse-translated to human computer games. Importantly, a long-standing tradition of lesion studies in rodents and an emerging field of lesion and functional imaging studies in humans 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, 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 ref. 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.

Here, we apply a decision-theoretic outlook on choice sequences in approach–avoidance conflicts 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 (DVs), specifically of a threat-related DV in the 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, and thus links approach–avoidance conflicts to the burgeoning literature on complex decision behaviour in ecological scenarios.

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. 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 pay-offs). For each mini-block of trials, called 'forest', energy was reset and foraging options varied. During the initial forest phase (3.5 s), participants were informed about their initial energy (2, 3 or 4 energy points) and the 2 possible foraging options in the current forest over 5 trials, called 'days'. The number of days was not depicted and varied according to an exponential distribution (for MRI design efficiency) but participants knew that their final pay-offs would depend on a subset of forests for which they would complete exactly 5 days. Triangles of 4 different sizes showed the current predator probability (0.1, 0.2, 0.3 or 0.4). Green rectangles of 4 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 − (probability of foraging gain)). After a fixation interval (3.5 s), the choice phase started (2.0 s): 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 (1.0 s), the outcome phase started (3.5 s). The choice phase started (2.0 s): one of the two foraging options in the current forest over 5 trials, called 'days'. The number of days past was not depicted on the screen.

The task was mathematically specified as a Markov decision process (MDP). 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 gain as well as further task variables in a mostly nonlinear fashion (Supplementary Fig. 1). We use the word ‘optimal’ here in the sense of ‘optimal under task instructions’ (that is, 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 DVs could potentially explain participants’ behaviour. To explain participants’ choices, we assessed an extensive set of potential DVs: (1) the optimal policy and (2–16) 15 possible DVs that are not optimal and are therefore regarded 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 heuristics 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 only for the current day (that is, if the time horizon was 1 day); two pseudo-optimal policies would be optimal if participants had to solely minimize the threats of either (11) starvation or of (12) predation. Four further DVs were included to address suggestions by anonymous reviewers. One of these DVs (13) was a pseudo-optimal policy that weighted the time horizon of the optimal policy.
Participants’ choices relied on predator probability and optimal policy. We 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 Fig. 2a for the metrics of the Bayesian information criterion (BIC) and Supplementary Table 1 for BIC and protected exceedance probabilities). Searching for the predictor that best explained exactly according to the distribution of the number of days that participants actually remained in a forest (that is, according to an exponential distribution with an average of 2.5, 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.Variables 1 to 16 were used in models of choice behaviour. On the basis of the two variables in the behavioural model (that is, on the basis of variables 1 and 2), we derived variables 17 to 19 and included these in RT and fMRI analyses.

| Variable | Description |
|----------|-------------|
| (1) Optimal policy | The optimal DV that takes into account all remaining time points (days), energy states and the transition probabilities between them. Therefore, the optimal policy cannot be directly inferred from information on the screen but necessitates rather complex computations. |
| (2) Predator probability (P(predictor)) | The probability of the predator attacking when the foraging option is chosen. This probability is depicted by the size of a black triangle on the screen and takes a value of 0.1, 0.2, 0.3 or 0.4. A predator attack leads to immediate death (that is, an energy state of 0). |
| (3) Probability of foraging gain (P(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 a value 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 0 to 4 and are depicted by the number of blue dots within the green rectangle. Waiting entails a sure loss of 1 energy point and unsuccessful foraging always entails a constant loss magnitude of 2 points (depicted as dark red dots). |
| (5) Continuous energy | Energy varies on a continuous scale from 1 to in steps of 1 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 1 is special because in that state waiting leads to sure death. The ‘binary energy’ variable therefore distinguishes between an energy state of 1 and higher 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 takes only one time step into account. |
| (8) Expected energy change | The difference between the current energy state and the expected energy state for foraging (see variable 7). |
| (9) Days past | The number of days (that is, time steps) already spent in a given forest. The maximum number of days within a forest is always 5. |
| (10) Pseudo-optimal: horizon-1 | This policy is optimal in the final time step (that is, when only 1 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 were not rewarded according to a full horizon of 5 days but were instead rewarded according to the horizon given by an exponential distribution with a mean of 2.5, which was implemented in the main experiment to enhance fMRI design efficiency. |
| (14) Past energy change | The difference in 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. The 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 before the current forest. |
| (17) Uncertainty of P(predator) | When the prescriptions of the employed heuristic policy (that is, of the predator probability) are closer to 0 (that is, waiting) or 1 (that is, foraging), choice uncertainty is lower than when the prescriptions lie in-between. Choice uncertainty is indexed by the derivative of the mean of the logistic functions for the predator probability (see Fig. 2c). |
| (18) Uncertainty of optimal policy | When the absolute value differences according to the optimal policy are large (that is, either clearly prescribing waiting or foraging), choice uncertainty is lower than when the absolute value differences are small (that is, the optimal policy is more or less indifferent). Choice uncertainty is indexed by the derivative of the mean of the logistic functions for the optimal policy (see Fig. 2d). |
| (19) Discrepancy in choice probabilities between P(predator) and optimal policy | In some cases, the heuristic policy of using the predator probability 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 Fig. 2c versus Fig. 2d). |

Variables 1 to 16 were used in models of choice behaviour. On the basis of the two variables in the behavioural model (that is, on the basis of variables 1 and 2), we derived variables 17 to 19 and included these in RT and fMRI analyses.
Fig. 2 | Models of choice data in the fMRI sample. **a**, Bayesian model comparisons show that predator probability was the best single predictor of participants’ choices. The plot depicts fixed-effects analyses using relative log-group Bayes factors, based on the BIC relative to the first model in the comparison set (here ‘optimal policy’). **b**, A model that additionally included the optimal policy best explained remaining variance in participants’ choices (log-group Bayes factors relative to the model including predator probability and optimal policy). **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 (that is, larger negative values). All error bars represent the s.e.m. 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 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 Fig. 1 for the relationships among the 16 DVs included in the models. See Supplementary Figs. 2 and 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 Fig. 4 depicts individual variability in the fMRI sample.
the remaining variance revealed that participants additionally relied on the optimal policy (Fig. 2b). The model including both predator probability and optimal policy outperformed the simpler model that included only 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 12 DVs motivated a priori (the additional 4 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, Fig. 2c; optimal policy, Fig. 2d) and in posterior predictive checks splitting data according to the variables not included in the model (Supplementary Fig. 2). Conversely, models with the other DVs did not capture the pattern of empirically observed choices (Supplementary Fig. 3). Individual differences in the best-fitting models are illustrated in Supplementary Fig. 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 Fig. 2 and Supplementary Tables 4–6). That is, predator probability decisively emerged as the best single 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 differed only 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). Owing 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 Fig. 1).

More importantly, confusion analyses with 2,000 simulations per model showed that models with 1 of the 16 considered DVs could be almost perfectly recovered from simulated data. This means that 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 1 simulated model was misclassified (Supplementary Table 8). Less than 10% misclassifications occurred for the last 3 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 1 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 3 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 the heuristic identified in the model comparisons corresponded to participants’ consciously accessible decisions.

RTs scaled with optimal policy and choice uncertainty of predator probability. We predicted that participants’ use of predator probability and 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 models (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 probability and optimal policy. In sum, RTs corroborate that predator probability and the optimal policy exert a joint influence on participants’ decision process on the same trials.

fMRI data could be related to DVs 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 and 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 dorsolateral prefrontal cortex 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 conflict tests18,26,27, predator probability was positively related to blood oxygen level-dependent (BOLD) signals in a cluster in the right anterior hippocampus extending into the neighbouring amygdala, as well as bilateral clusters in the dorsolateral prefrontal cortex (DLPFC; Fig. 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 (Fig. 3b). This cluster overlapped with the cluster identified for predator probability per se (see Fig. 4a for the overlap of the two functional clusters and for information on the overlap with anatomical hippocampus and amygdala masks; see Fig. 4b for the relation of parameter estimates to predator probability). Uncertainty of the predator probability was also negatively associated with BOLD signals in the bilateral DLPFC and in the 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 Fig. 5).

Taken together, both the predator probability and its associated choice uncertainty scaled on a trial-by-trial basis with BOLD signals 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; Fig. 3c and Table 3), extending into the supplementary motor area (SMA). The same parametric contrast revealed a positive relation of the optimal policy with the anterior cingulate cortex (ACC), bilateral IFG (extending into the insula) and bilateral thalamus. The choice uncertainty of the optimal policy scaled positively with BOLD signals in DMPFC and DLPFC regions, as well as in the IFG and insula (Fig. 3d). The same metric scaled negatively with activity in the posterior cingulate cortex. In the DMPFC and IFG (extending into the insula), we found overlaps between clusters elicited by the optimal policy per se and by the choice uncertainty of the optimal policy (Supplementary Fig. 5). Furthermore, BOLD signals in the thalamus showed a positive association with the discrepancy in choice probabilities between predator probability and 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 (that is, energy state at outcome minus energy state at choice) was robustly associated with neural activity in a well-established reward network (that is, the bilateral striatum, ventral medial prefrontal cortex and posterior cingulate cortex; Supplementary Fig. 6 and Supplementary Table 13). 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 strengthened and illustrated 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 the bilateral hippocampus but failed to reach whole-brain family-wise error (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 the bilateral striatum co-varied negatively with individual parameter estimates capturing the degree to which predator probability influenced participants’ choices (Supplementary Fig. 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 the respective DV (Supplementary Fig. 5). Following the suggestions
Fig. 3 | fMRI results during the choice phase. a, Predator probability showed a positive relation within a cluster spanning the right anterior hippocampus and the amygdala, as well as within the bilateral DLPFC. b, The choice uncertainty according to the predator probability showed a negative relation in the right anterior hippocampus, extending into the amygdala as well as the bilateral DLPFC and the right lateral IFG. c, The optimal policy showed a positive relation within DMPFC, extending into pre-SMA and ACC, as well as into the thalamus and other regions. d, The choice uncertainty according to the optimal policy showed a positive relation in the DMPFC extending into the pre-SMA. Number of participants, n = 24. fMRI clusters were overlaid on a group average T1-weighted image in MNI space; clusters are whole-brain 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 Fig. 4 for the overlap of the hippocampus/amygdala clusters depicted in a and b. Supplementary Fig. 5 visualizes the overlap of all clusters. See Supplementary Table 12 and Supplementary Fig. 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 Fig. 8 present the results of covariate analyses testing for inter-individual differences. Supplementary Table 16 and Supplementary Figs. 8 and 9 present non-independent ROI analyses. See Supplementary Table 17 for contrasts testing for interaction effects. Supplementary Fig. 10 visualizes the overlap between clusters from the current study and clusters from a related previous study2.

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 predation11,21 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 threats1,8. Crucially, our results associate the neural controller of the heuristic policy with structures often implicated in approach–avoidance conflicts in both rodents and humans11–16,18,21. That is, the anterior hippocampus and the amygdala related to predator probability as well as to the uncertainty of using this heuristic 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-making1,3,6.

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 layouts18,26,27. 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 predation1. This may indicate that

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 probably suggests that these regions were influenced by more than one behavioural variable (Supplementary Figs. 8 and 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 (Fig. 4 and Supplementary Fig. 5), the clusters in the hippocampus and the amygdala as well as the clusters in the bilateral DLPFC were positively related to predator probability per se and negatively related to the associated choice uncertainty (Supplementary Table 16). Similarly, the DMPFC and right IFG (extending into the 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 in separate GLMs (that is, 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).
| Side          | Peak voxel MNI coordinates (mm) | Cluster size (voxels) | Peak t score |
|--------------|---------------------------------|-----------------------|--------------|
| (1) P(predator): positive (trial-by-trial relation with higher numbers indicating higher probability of predator attacking) |
| DLPFC        | R                               | 30 36 48              | 1,325        | 6.69         |
| Medial occipital cortex | L                             | −14 −74 −9          | 1,112        | 6.17         |
| Anterior hippocampus extending into the 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 higher value of the foraging option versus the waiting option according to the optimal policy) |
| Thalamus     | L                               | −8 −21 6             | 250          | 7.92         |
| IFG extending into the insula | R                             | 32 23 5             | 950          | 7.58         |
| ACC          | 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 DMPFC extending into the SMA | L/R                          | 8 23 53             | 1,393        | 5.57         |
| IFG extending into the 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 the DMPFC | R                             | 14 38 54            | 1,977        | 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 the 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 and DMPFC | R                             | 17 12 57            | 2,484        | 6.96         |
| Insula       | R                               | 21 23 −5            | 707          | 6.37         |
| DLPFC        | R                               | 39 35 41            | 935          | 6.16         |
| DLPFC extending into the 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         |
| (5) Discrepancy between P(predator) and 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) and optimal policy: negative |
| None         |                                 |                      |              |

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, $n=24$. 
participants are able to select a particularly appropriate heuristic for the task at hand (that is, a heuristic that combines computational simplicity with near-optimal approximations). Given several theoretical accounts from multiple fields arguing for adaptive heuristic decision-making [14,15], it is an interesting avenue for future research to investigate how participants select and switch between different heuristics under varying biologically inspired decision tasks [16].

Neurally, different heuristics are likely to be implemented by different controllers. 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]. Our fMRI results cannot delineate subtle differences in the roles of the 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 [17]. 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 [18].

With respect to work on humans, our findings relating the anterior hippocampus and also the amygdala to predator probability corroborate recent studies implicating these structures in different types of approach–avoidance conflicts [19–24]. 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 [25]. The relation between predator probability and the hippocampus in our rather strategic task fits particularly well with the recently demonstrated involvement of the hippocampus during strategic decisions to escape from a slow-attacking virtual predator [26]. Overall, the findings linking the hippocampus and amygdala with risk 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 [27].

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 or inferences about abstract variables and sensory stimuli [28–30]. Here, we were not interested in outcome uncertainty but rather in choice uncertainty (that is, 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 [31].

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 [33]. Previous research has not been able to address whether humans are capable of using elaborate decision policies for surmounting approach–avoidance conflicts. We could do so because our sequential decision-making task was formulated in the precise mathematical framework of an MDP (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 ACC as well as the 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 [34–36]. The implication of the DMPFC and the ACC might also point to further converging neural processes across
specifies two overlapping regions of the hippocampus and amygdala. These studies found four overlapping areas of the hippocampus and amygdala—along with the associated choice uncertainty. These findings argue against a monoclonal 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. 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 >1 mm during fMRI, one due to an incidental medical finding revealed by MRI, and three who behaved almost deterministically (that is, they selected one of the two choice options in more than 0.85 of the current sample in Hamburg). The final sample comprised 23 participants (15 female; age = 25.9 ± 3.5 years). Due to time constraints, one participant performed only eight out of ten sessions. Participants received a show-up fee of EUR12 plus a variable amount.

The study was conducted in accord with the Declaration of Helsinki and approved by the ethics committees of the Kantonale Ethikkommission Zürich, KEK-ZH-Nr. 2013-0328; behavioural sample: Ethikkommission der Ärztekammer Hamburg, PV 5746. All participants gave written informed consent using a form approved by the ethics committees.

Sequential decision-making task. Instructions and task. See Fig. 1 for an overview of the task set-up. Participants performed only nine out of ten sessions. The task was presented as virtual foraging under the dual threats of predation and of starvation. To familiarize themselves with the task, participants performed 2 training sessions: a first short training session of 4 forests with 5 days each (after which participants could ask questions) and a second longer training session of 24 forests with 5 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 condensed (that is, 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 fMRI 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 days in ten for exploratory analyses, and 5 days in ten for the final analyses. 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 CHF1.50 (fMRI sample) or EUR1 (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’ and of the Need for Cognition scale.

Mathematical framework and optimal policy. We modelled the task as an MDP. 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; that is, 6 energy states (0–5 energy points) × 2 (weather types).

(2) Actions. Foraging or waiting.

(3) Transition matrix. This is constructed from the probabilities of predator attack and the probabilities of foraging gain, gain 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 (that is, day five) and calculated the values of the two choice options (that is, 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 it is chosen. Conversely, 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 (that is, 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. This procedure was repeated until arriving at the first time step (that is, day one).

The optimal policy per sea specifies the action to choose and thus does not allow for variability in the decision process (that is, 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 considers just a finite time horizon of one time step. A second pseudo-optimal policy neglects the transitions according to the predator probability (that is, 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 1 to 5 days according to the number of times these horizons actually occurred in the main experiment (that is, the number of days according to an exponential distribution with a mean of 2.5, 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 (Fig. 2c,d). These derivatives capture the intuition that a small deviation in the DV (for example, due to perceptual or computational error) has a larger impact on the resulting choice at the inflection point of the logistic function (that is, at the point at which participants are indifferent between the two choice options) than at the ranges of the DV 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 RT data. Participants received a total of 400 trials (that is, days except for 2 participants in the fMRI who received only 360 trials due to equipment malfunction and 1 participant in the behavioral sample who received only 320 trials due to time constraints). On average, participants were ‘alive’—and were thus able to make decisions—in 354.4 ± 16.1 trials, mean ± s.d., (fMRI sample) and 374.4 ± 21.9 trials (behavioral sample). Of these, they failed to answer in 8.4 ± 8.8 trials (fMRI sample) and 5.4 ± 8.3 trials (behavioral sample), which left 346.0 ± 17.6 trials (fMRI sample) and 349.0 ± 27.6 trials (behavioral sample) for analyses.

To explain participants’ choices, that is, their probability of choosing the foraging option, \( p_{\text{optim}} \), we used logistic regression models (implemented in the MATLAB function mnrfit) of the following generic form (where parameter estimates are denoted by \( \beta \)):

\[
p_{\text{optim}} = 1/(1 + \exp(-DV))
\]

with the following form of the DV:

\[
DV = \beta_0 + \beta_1 \times \text{predactor}
\]

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

\[
DV = \beta_0 + \beta_1 \times \text{predactor}_1 + \beta_2 \times \text{predactor}_2
\]

We also tested interaction models:

\[
DV = \beta_0 + \beta_1 \times \text{predactor}_1 + \beta_2 \times \text{predactor}_2 + \beta_{12} \times \text{predactor}_1 \times \text{predactor}_2
\]

For each model we approximated model evidence by calculating the BIC, which penalizes model complexity:

\[
\text{BIC} = 2 \times \text{ln} (L) - k \times \text{ln}(n)
\]

where SDR is the sum of deviance residuals (that is, 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 assumes 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. For the Bayesian model selection procedure, BIC values were multiplied by −0.5 to scale values with respect to the appropriate conventions.

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/lme4/index.html). 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). LLDS 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 3T (Philips Achieva) MR scanner using a 32-channel head coil. Functional images were recorded using a T2*-weighted echo-planar imaging (EPI) sequence (repetition time (TR) 2.1 s; echo time (TE) 30 ms; flip angle 90°). A total of 30 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 10 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 349.1 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°). Functional images were acquired using a T1-weighted scan (field of view 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. Preprocessing of EPI data included rigid-body realignment to correct for head movement, unwarping and slice time correction. EPI images were then co-registered to the individual’s T1-weighted image using the formation of 12-parameter affine transformation. The two normal maps were transformed into the Montreal Neurological Institute (MINI) T1 reference brain template using the extended unified segmentation algorithm in SPM12. 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 haemodynamic response function and low-frequency drifts were excluded using a high-pass filter with a 128 s cutoff.

In the GLM the three distinct phases of the task (forest, choice and outcome phases; see Fig. 1) were entered as events with a duration of 0 s (that is, 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 Fig. 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 and 4) the associated choice uncertainties; (5) the difference between two metrics and (6) log-transformed RTs. We report analyses in which parametric modulators competed for variance (that is, 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 model except for some numerical differences due to rounding. Due to collinearity 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 (that is, energy state at outcome minus energy state at choice).

In a second GLM, participants’ choices were included as an additional parametric modulator (that is, as a binary modulator coding for waiting or foraging; see Supplementary Table 13). In a tertiary GLM, choice uncertainties and discrepancy were calculated from the independent value differences (see Supplementary Table 14). For follow-up ROI analyses, we set up a GLM with four separate onset regressors for the four levels of predator probability (Fig. 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 (that is, 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 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. ROI analyses were conducted using the toolboxes rFplot and marsbar.
Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

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/3046/).

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

Received: 3 July 2018; Accepted: 3 April 2019; Published online: 20 May 2019

References
1. Bach, D. R. & Dayan, P. Algorithms for survival: a comparative perspective on emotions. Nat. Rev. Neurosci. 18, 311–319 (2017).
2. Korn, C. W. & Bach, D. R. Heuristic and optimal policy computations in the human brain during sequential decision-making. Nat. Commun. 9, 325 (2018).
3. Korn, C. W. & Bach, D. R. Maintaining homeostasis by decision-making. PLoS Comput. Biol. 11, e1004301 (2015).
4. Huys, Q. J. M. et al. Interplay of approximate planning strategies. Proc. Natl Acad. Sci. USA 112, 3098–3103 (2015).
5. Huys, Q. J. M. et al. Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees. PLoS Comput. Biol. 8, e1002410 (2012).
6. Keramati, M., Smittenaar, P., Dolan, R. J. & Dayan, P. Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. Proc. Natl Acad. Sci. USA 113, 12868–12873 (2016).
7. Keramati, M., Duzel, A. & Piray, P. Speed/accuracy trade-off between the habitual and the goal-directed processes. PLoS Comput. Biol. 7, e1002055 (2011).
8. LeDoux, J. & Daw, N. D. Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. Nat. Rev. Neurosci. 19, 269–282 (2018).
9. Mobbs, D., Trimmer, P. C., Blumstein, D. T. & Dayan, P. Foraging for foundations in decision neuroscience: insights from ethology. Nat. Rev. Neurosci. 19, 419–427 (2018).
10. Mobbs, D. The ethological deconstruction of fear(s). Curr. Opin. Behav. Sci. 24, 32–37 (2018).
11. Griebel, G. & Holmes, A. 50 years of hurdles and hope in anxiolytic drug development. Br. J. Pharm. 162, 69–74 (2017).
12. Griebel, G. & Holmes, A. The ethological deconstruction of fear(s). J. Neurosci. 37, 1249–1254 (2018).
13. Griebel, G. & Holmes, A. 50 years of hurdles and hope in anxiolytic drug development. Br. J. Pharm. 162, 69–74 (2017).
14. Kolling, N., Behrens, T. E. J., Mars, R. B. & Rushworth, M. F. S. Neural mechanisms of foraging. Science 336, 95–98 (2012).
15. Kolling, N., Wittmann, M. & Rushworth, M. F. S. Multiple neural mechanisms of decision making and their competition under changing risk pressure. Neuron 81, 1190–1202 (2014).
16. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. Nat. Neurosci. 14, 933–939 (2011).
17. Mata, R., Wilke, A. & Czienskowski, U. Foraging across the life span: is there a reduction in exploration with aging? Front. Neurosci. 7, 53 (2013).
18. ShenHAV, A., Straccia, M. A., Cohen, J. D. & Botvinick, M. M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat. Neurosci. 17, 1249–1254 (2014).
19. Constantino, S. M. & Daw, N. D. Learning the opportunity cost of time in a patch-foraging task. Cogn. Affect. Behav. Neurosci. 15, 837–853 (2015).
20. Pearson, J. M., Watson, K. K. & Platt, M. L. Decision making: the neuroethological turn. Neuron 82, 950–965 (2014).
21. Giogreneraet, G. & Gaisser, G. M. Heuristic decision making. Annu. Rev. Psychol. 62, 451–482 (2011).
22. Gu, X. & FitzGerald, T. H. B. Interceptive inference: homeostasis and decision-making. Trends Cogn. Sci. 18, 269–270 (2014).
23. Fawcett, T. W. et al. The evolution of decision rules in complex environments. Trends Cogn. Sci. 18, 153–161 (2014).
24. Dayan, P. Rationalizability of irrational choices. Top. Cogn. Sci. 6, 204–228 (2014).
25. Dayan, P. Rationalizability of irrational choices. Top. Cogn. Sci. 6, 204–228 (2014).
26. Toyota, P., Fadok, J. P. & Läthi, A. Neuronal circuits for fear and anxiety. Nat. Rev. Neurosci. 16, 317–331 (2015).
27. Jimenez, J. et al. Anxiety cells in a hippocampal–hypothalamic circuit. Neuron 97, 670–683.e7 (2018).
28. Blanchard, D. C. Translating dynamic defense patterns from rodents to people. Neurosci. Biobehav. Rev. 76, 22–28 (2017).
29. Payzan-LeNestour, E., Dunne, S., Bossaerts, P. & O’Doherty, J. The neural representation of unexpected uncertainty during value-based decision making. Neuron 79, 191–213 (2013).
30. Rigoli, F., Michely, J., Friston, K. J. & Dolan, R. J. The role of the hippocampus in weighting expectations during inference under uncertainty. Cortex 115, 1–14 (2019).
31. Harrison, L. M., Duggins, A. K. & Friston, K. J. Encoding uncertainty in the nervous system. Neuron 97, 670–683.e7 (2018).
32. Blanchard, D. C. Translating dynamic defense patterns from rodents to people. Neurosci. Biobehav. Rev. 76, 22–28 (2017).
33. Rigoli, F., Michely, J., Friston, K. J. & Dolan, R. J. The role of the hippocampus in weighting expectations during inference under uncertainty. Cortex 115, 1–14 (2019).
34. Stranges, B. A., Duggins, A. K., Penny, W., Dolan, R. J. & Friston, K. J. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? Neural Netw. 18, 225–230 (2005).
35. Lee, S. W., Shimojo, S. & O’Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. Neuron 81, 667–699 (2014).
Acknowledgements
We thank G. Castegnetti, S. Khemka, M. Staib, A. Tzovara and C. 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). C.W.K. 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.

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.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0603-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to C.W.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

### Statistical parameters

| n/a | Confirmed |
|-----|-----------|
| ☒   | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☒   | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒   | Only common tests should be described solely by name; describe more complex techniques in the Methods section. |
| ☒   | A description of all covariates tested |
| ☒   | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☒   | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒   | For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted |
| ☒   | Give P values as exact values whenever suitable. |
| ☒   | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒   | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒   | Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated |
| ☒   | Clearly defined error bars |
| ☒   | State explicitly what error bars represent (e.g. SD, SE, CI) |

Our web collection on statistics for biologists may be useful.

### Software and code

Policy information about availability of computer code

| Data collection | The behavioral task was presented to human participants using the MATLAB toolbox Cogent (www.vislab.ucl.ac.uk). |
| Data analysis   | Behavioral data was analyzed using custom code in MATLAB and especially using logistic regression models (implemented in the MATLAB function mnrfit). We used the log-group Bayes factors based on Bayesian Information Criterion (BIC) and the Bayesian Model Selection (BMS) procedure implemented in the MATLAB toolbox SPM12 (http://www.fil.ion.ucl.ac.uk/spm/) to calculate protected exceedance probabilities. We analyzed 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) (Baayen et al., 2008). All MRI analyses were performed in SPM12. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.
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/).

**Field-specific reporting**

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- ☐ Life sciences
- ☑ Behavioural & social sciences
- ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

### Study description

Participants performed a decision-making task in the MR scanner. Data are quantitative, i.e., chosen options, RTs, BOLD signals. An additional sample was tested behaviorally.

### Research sample

The study comprised students from local universities who were recruited via mailing lists. The final fMRI sample comprised 24 participants (11 female; age = 25.0 ± 3.9 years). The research sample is thus comparable to many fMRI studies. Exclusion criteria for participation were the standard MR incompatibility criteria. The final sample behavioral comprised 23 participants (15 female; age = 25.9 ± 3.5 years).

### Sampling strategy

Sampling was by random order of reply to our recruitment. No explicit sample size calculation was performed. Sample size was based on standards in the field and on experience from previous studies.

### Data collection

MRI data were acquired on a 3 T (Philips Achieva, Best, The Netherlands) MR scanner. Behavioral data were collected via button boxes. Besides participants and researchers no other person was present during data collection. The researchers were not blind to the study hypothesis. Conditions did not differ between participants such that all participants received the same instructions and treatment.

### Timing

Data were collected from September to November 2015 for the fMRI sample and from November 2018 to January 2019 for the behavioral sample.

### Data exclusions

FMRI sample: 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. Behavioral sample: 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.

### Non-participation

None of the invited participants dropped out or declined participation.

### Randomization

Conditions did not differ between participants. Within participants, conditions were randomly ordered.

### Reporting for specific materials, systems and methods
### Human research participants

**Policy information about studies involving human research participants**

- **Population characteristics**: The final fMRI sample comprised 24 participants (11 female; age = 25.0 ± 3.9 years). The final behavioral sample comprised 23 participants (15 female; age = 25.9 ± 3.5 years).

- **Recruitment**:
  - FMRI sample: We recruited 29 participants via mailing lists of local universities.
  - Behavioral sample: We recruited 26 participants via a local online platform.

### Magnetic resonance imaging

#### Experimental design

- **Design type**: decision-making task, event-related
- **Design specifications**: 10 scanning sessions per participant; 40 trials per scanning session; depending on condition trials were either 11 or 4 sec long; between trials there was a variable fixation interval (ranging from 0.5 to 3.8 sec).
- **Behavioral performance measures**: button presses and RTs were recorded; we excluded participants who behaved almost deterministically, i.e. selected one of the two choice options in more than 0.9 of the retained trials.

#### Acquisition

- **Imaging type(s)**: functional
- **Field strength**: 3
- **Sequence & imaging parameters**: 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°).
- **Area of acquisition**: whole brain
- **Diffusion MRI**:
  - Used: No
- **Preprocessing**:
  - **Preprocessing software**: SPM12. The FieldMap toolbox was used to correct for geometric distortions caused by susceptibility-induced field inhomogeneities. Normalized images were smoothed with an isotropic 8 mm full width at half maximum Gaussian kernel.
  - **Normalization**: EPI images were 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 (Ashburner and Friston, 2005).
  - **Normalization template**: MNI305
  - **Noise and artifact removal**: The six motion correction parameters estimated from the realignment procedure were entered as covariates of no interest.
  - **Volume censoring**: None

#### Statistical modeling & inference

- **Model type and settings**: Mass-univariate. 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). We were mostly interested in the choice phase and ran a primary GLM with a combination of variables that emerged in our analyses of behavioral and RT data.
Effect(s) tested
We performed second-level one-sample t-tests (one-sided) on contrast images from all participants. Our focus was on the parametric regressors that were used to model the choice phase.

Specify type of analysis: □ Whole brain  □ ROI-based  □ Both

Statistic type for inference
(See Eklund et al. 2016)
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 (Eklund et al., 2016).

Correction
FWE correction using the SPM random field theory based approach.

Models & analysis
n/a
☑ Functional and/or effective connectivity
☑ Graph analysis
☑ Multivariate modeling or predictive analysis