



Published in final edited form as:

Nat Neurosci. 2014 September ; 17(9): 1249–1254. doi:10.1038/nn.3771.

Anterior Cingulate Engagement in a Foraging Context Reflects Choice Difficulty, Not Foraging Value

Amitai Shenhav¹, Mark A. Straccia¹, Jonathan D. Cohen¹, and Matthew M. Botvinick¹

¹Princeton Neuroscience Institute and Department of Psychology, Princeton University, Princeton, NJ 08540

Abstract

Previous theories predict that human dorsal anterior cingulate (dACC) should respond to decision difficulty. An alternative theory has been recently advanced which proposes that dACC evolved to represent the value of “non-default,” foraging behavior, calling into question its role in choice difficulty. However, this new theory does not take into account that choosing whether or not to pursue foraging-like behavior can also be more difficult than simply resorting to a “default.” The results of two neuroimaging experiments show that dACC is only associated with foraging value when foraging value is confounded with choice difficulty; when the two are dissociated, dACC engagement is only explained by choice difficulty, and not the value of foraging. In addition to refuting this new theory, our studies help to formalize a fundamental connection between choice difficulty and foraging-like decisions, while also prescribing a solution for a common pitfall in studies of reward-based decision making.

The dorsal anterior cingulate cortex (dACC) currently stands out as one of the most extensively studied regions of the brain, and yet its basic functions are still a matter of intensive debate^{1–7}. Historically, functions attributed to this region have included the encoding of pain⁵, surprise^{3, 8}, value⁹, and level of cognitive demand^{5, 10, 11}, including the difficulty posed by conflict between competing choices^{12–16}.

A recent study by Kolling, Behrens, Mars, and Rushworth¹⁷ (KBMR) presented a significant and intriguing detour from previous approaches to understanding the function of this region. The authors echo previous proposals that the dACC is responsive to the value of choice options. However, they propose that this is restricted to the value of diverging from one’s default behavior in a given context. They suggest that this function has its evolutionary roots in the encoding of the overall value of *foraging* for food in a new patch

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 and requests for materials should be addressed to A.S. (ashenhav@princeton.edu).

Author contributions: A.S., J.D.C., and M.M.B. designed the experiments; A.S. and M.A.S. performed the experiments; A.S. analyzed the data; and A.S., J.D.C., and M.M.B. wrote the paper.

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

The authors declare no competing financial interests.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

rather than continuing to *engage* the current food patch. To test this, they designed a task to model foraging decisions. KBMR showed that, in this task, dACC activity was positively associated with the value of “foraging” for better rewards and negatively associated with the value of “engaging” currently available ones (the default behavior). Their conclusion, that dACC is involved in foraging decisions per se, has already had a significant impact on theorizing regarding the function of this region^{1, 18} and has generated several high-profile follow-up studies that reach similar conclusions^{19–21}.

In the present work, we challenge KBMR’s interpretation of their findings and present strong evidence that dACC’s role in foraging-like decisions is instead connected with decision difficulty.

Although it has not been widely remarked, standard theories of foraging imply an intimate relationship between foraging and decision difficulty. Figure 1 illustrates the dynamics of foraging, focusing on two quantities: the rate of intake within a patch (blue line), and the value of foraging, i.e., leaving for another patch (green line). According to optimal foraging theory^{22, 23, 24}, the best moment to leave a patch is when these two values coincide. If (like KBMR) we view foraging choices as involving comparisons between two values or utilities, it appears that the optimal moment to forage is precisely the moment at which the value-based decision becomes most difficult, that is, the moment at which the values to be compared are most similar.

A more careful look at the pattern described above (Fig. 1) forces a reconsideration of KBMR’s findings. Note in particular the relationship between foraging value (green) and decision difficulty (proximity of green and blue). These two are closely linked: As foraging value rises, so does decision difficulty. This correlation raises a serious concern, in connection with the theory advanced by KBMR and related work. Specifically, it suggests that foraging value and decision difficulty might be confounded in the experiments that motivate the theory.

We show here, based on two fMRI experiments, not only that these two factors have been confounded in previous studies, but that when they are adequately dissociated, dACC activity is found clearly to track choice difficulty rather than foraging value. We begin with a replication of the KBMR study, revealing a fundamental problem with the measure of choice difficulty used in that study, and introducing a more principled measure based on observed choice behavior. In a second experiment, we take advantage of features of the KBMR task that distinguish it from standard foraging tasks to deconfound choice difficulty from foraging value. When we do this, we obtain precisely the opposite result from the one they reported: Anterior cingulate activity in foraging tasks tracks choice difficulty rather than foraging value. The set of results we report highlight (1) a fundamental point about the structure of foraging-like tasks, and the potential role of dACC in the performance of such tasks, and (2) a key methodological consideration in studies of value-based choice.

Results

In Experiment 1 we scanned 15 human participants while they performed KBMR's original foraging task. Each trial involved two stages of decision-making (Fig. 2a). In Stage 1, the participant was offered a pair of potential rewards (the *engage* set) and a set of six alternative possible rewards (the *forage* set), all presented as abstract symbols designating numerical points that could be earned. The participant could choose to proceed directly to Stage 2 (engage option) or first swap the current engage pair for a new pair randomly selected from the forage set as many times as they wished (forage option). Each swap generated a new forage set but incurred a cost (a designated number of lost points, as well as a time delay). Choosing to engage advanced the trial to Stage 2, at which time a probability was paired with each of the two options in the engage set, a choice was made between the two reward-probability pairs, and the chosen reward was received — based on the outcome of a random draw — with the corresponding probability.

KBMR reported three key findings from this task. First, in Stage 1 they found a strong bias to engage. Second, they found that dACC tracked the degree to which the points offered for each option favored the foraging option — that is, increased its relative value (RV_{forage}) as compared to the engage option. Third, in Stage 2, they found no prepotent bias between the two options (left vs. right), and found that dACC activity increased as the expected value of the two engage options became more similar. We replicated all of these findings (Fig. 3a–b; Supplementary Fig. 1). KBMR point out that the third finding is consistent with accounts that predict that dACC activity should track the difficulty of a given choice^{4, 14}, but argue that the second finding cannot be explained in these terms. Critically, they argue that a difficulty account of dACC should predict greatest activity when choices to forage and engage were equivalent in point value (that is, at $RV_{forage} = 0$). But this was not so; rather, dACC activity continued to increase past this point, as forage value continued to increase. KBMR interpreted this as evidence that dACC activity coded the value of the foraging option and not choice difficulty. Furthermore, noting the strong overall bias to engage, they argued that the default action was to engage and that the association of dACC activity with the value of foraging was consistent with a role in encoding the value of the non-default action. For Stage 2, they postulated that the unchosen option was *prima facie* the non-default, and thus dACC activity increased as its value increased (i.e., it became more similar to the chosen one).

KBMR's task confounds foraging value and choice difficulty

On the surface, KBMR's evidence for a foraging account of dACC may seem persuasive. However they depend upon a problematic assumption, namely that difficulty was greatest when the points offered for the two options were most similar (that is, when RV_{forage} was closest to zero). The validity of this assumption can be checked against the empirically observed choice behavior, by plotting RV_{forage} against the observed frequency of choices to forage. If KBMR's assumption is correct, then the point on the curve at which RV_{forage} equals zero should coincide with the point at which the forage and engage options each had a 50% likelihood of being selected, the *empirical indifference point*. It should similarly coincide with the point at which decisions take the longest, reflecting a maximum in choice

difficulty. For the unbiased choices made in Stage 2, we see that both conditions hold (Fig. 2c, black curves). However, in Stage 1, neither is the case: KBMR assumed that choice difficulty would reach its maximum at $RV_{forage} = 0$. Our results showed that both the empirical indifference point (Fig. 2b, left) and the point of longest response times (RTs) (Fig. 2b, right) aligned at a similar RV_{forage} value, but that this point was substantially to the right of $RV_{forage} = 0$.

Given these observations, a more appropriate measure of choice difficulty would adjust the inferred value of the options such that RV_{forage} is shifted to the right, and centered on each participant's empirical indifference point. To identify these empirical indifference points we fit a standard model of decision making^{4, 25–29} to each participant's choice and RT data (red curves in Fig. 2b–c; see Methods and Supplementary Fig. 2). The best fit of the model to these behavioral data required setting the indifference point, on average, at $RV_{forage} = 1.65$, significantly to the right of 0 (SE = 0.26, $t(14) = 6.5$, $p = 1.4 \times 10^{-5}$; Fig. 2b), and consistent with the observed bias to engage. We use each participant's estimated shift from this zero-point to generate estimates of relative foraging value that better match their empirical choice behavior. We refer to this corrected estimate as $RV_{forage-C}$.

While these analyses reveal a problem with KBMR's index of difficulty, they also uncover a more serious concern: When a more appropriate index of difficulty is applied (based on $RV_{forage-C}$), it becomes evident that KBMR's experimental design *confounded* foraging value and choice difficulty (Fig. 4a). To see this, consider the range of foraging choices used in their study. As shown in Figure 2b, this range is reasonably evenly distributed around the zero point for RV_{forage} , but not for $RV_{forage-C}$. Critically, the vast majority of choices fell to the left or near participants' actual choice indifference points, resulting in a high correlation between foraging value and an estimate of value similarity (choice difficulty) based on the absolute magnitude of $RV_{forage-C}$ ($-|RV_{forage-C}|$; average Spearman's $\rho = 0.80$, $t(14) = 14.1$, $p = 1.2 \times 10^{-9}$). Accordingly, when we regress BOLD activity on this measure of difficulty that accords more directly with observed behavior, we identify the same regions of dACC as KBMR found to track foraging value in Stage 1 and the value of the non-default (i.e., unchosen) option in Stage 2 (Fig. 3c and 5a–b).

Difficulty explains dACC activity better than forage value

KBMR's original analyses thus appear both to incorrectly model choice difficulty, and to confound it with foraging value. As suggested above, the latter is a consequence of the range of choices tested in their study. Critically, in this range both interpretations of the findings — in terms of foraging or difficulty — make the same prediction: activity of dACC should increase as the value of foraging increases and the pair of options approaches the indifference point (Fig. 4a). However, the two theories make different predictions as the value of foraging increases further, and begins to strongly favor selection of that option. The foraging theory predicts that dACC activity should continue to increase (or perhaps asymptote) in this range, as foraging continues to increase in value. In contrast, a difficulty-based account predicts that dACC should *decrease* past the indifference point, as the value of foraging more decisively exceeds that of engagement and thus choices become easier. In other words, the foraging account predicts a monotonic relationship between foraging value

and dACC activity, whereas the difficulty account predicts a non-monotonic relationship, with activity maximal at the point of indifference and dropping off as one option or the other becomes more clearly preferred and choices for that option are correspondingly more probable.

Experiment 2 (N=14) tested these predictions. We used a modified version of KBMR's task that maintained an engage bias ($M_{\text{indifference_point}} = 5.0$, $SE = 1.3$, $t(13) = 3.9$, $p = 0.0018$; see also Supplementary Fig. 1) but included choice sets that spanned a wider range, from ones that strongly favored the engage option to others that strongly favored the forage option (Fig. 4a–b; see Methods). When testing a range of foraging values comparable to Experiment 1, we again find a linear relationship between foraging value and dACC activity (Supplementary Fig. 3). However, as we test value ranges that increasingly favor the forage option (using a sliding windowed analysis) we see that dACC's relationship with foraging value becomes less positive and then reverses such that, at the upper end of foraging values, dACC has a significant *negative* correlation with foraging value (see also Supplementary Fig. 8). The results of Experiment 2 thus clearly demonstrated a non-monotonic relationship between foraging value and dACC activity, whereby activity was least when the choice options strongly favored either engaging *or* foraging, and greatest when they were equiprobable (Fig. 5c). Furthermore, as in Experiment 1, we found that choice difficulty explained dACC activity at both stages of the task (Fig. 5d). Finally, because choice difficulty and foraging value were orthogonal in this task ($M_{\rho} = 0.07$, $t(13) = 0.8$, $p = 0.45$), we were also able to directly compare the ability of each to predict dACC activity. We did so by entering them into the same general linear model, and found a significant effect of difficulty ($t(13) = 2.8$, $p = 0.014$) but not foraging value ($t(13) = 0.47$, $p = 0.64$) (Supplementary Fig. 4a). Two additional tests confirmed that dACC activity in this study was in fact better accounted for by difficulty than foraging value. First, a direct contrast of the regressors in the aforementioned GLM showed a significantly greater average parameter estimate for difficulty than foraging value (paired $t(13) = 2.2$, $p = 0.046$; see also Supplementary Fig. 4b). Second, we performed a Bayesian model comparison over the separate GLMs that accounted only for foraging value or only for choice difficulty, and found that the difficulty model was favored across dACC (Supplementary Fig. 4c), including within our *a priori* ROI (Supplementary Fig. 4d). These findings weigh heavily in favor of choice difficulty and against foraging value as an account of responses in dACC.

In keeping with analyses performed in KBMR's study, the analyses above assume that choice difficulty is a simple linear function of value similarity. Further analyses show that our findings hold for alternate formulations of difficulty that instead focus on the relative likelihood of choosing one option or another, based on either the decision model described above (the drift diffusion model²⁵; Supplementary Fig. 5b), or a simpler model of the decision process (i.e., a logistic regression; Supplementary Fig. 5c). These measures of difficulty have the benefit of accounting for nonlinearities in choice behavior (Fig. 2b–c and 4b), including the tendency for choice probabilities to asymptote beyond a certain point on the relative value scale. Of course, it is worth noting that similar activations are found when simply regressing dACC activity on RT, which is assumed to provide a “model-free” estimate of difficulty for two-alternate forced-choice tasks like this one. However, because

RT is a *noisy* estimate of difficulty we tested for and found a significant contribution of our earlier model-based estimate of difficulty ($-|RV_{forage-cl}$) to dACC activity even after removing variance accounted for by RT ($t_{Exp1}(14) = 4.0, p = 0.0013, t_{Exp2}(13) = 2.4, p = 0.034$). Thus, despite the close relationship between choice difficulty and RT on this task (as in many), we were still able to rule out a simple “time-on-task” (i.e., purely RT-based^{30, 31}) account of the choice difficulty effects we observed in dACC.

Consistent with the inherent relationship between foraging value and choice difficulty in standard foraging settings (Fig. 1), the potential for confounding these two variables lurks in any study that attempts to link dACC to foraging/non-default valuation. This appears to be the case for three prominent neuroimaging studies that have been argued to support KBMR’s foraging account. Mobbs and colleagues²¹ had participants perform an analog to a dynamic patch-leaving task, and showed that dACC activity increased as conditions favored leaving the patch (more competition and/or less reward for current resources) and decreased as conditions favored staying. As illustrated in Figure 1, to the extent that participant behavior in any way approximated optimal foraging, dACC activity was simultaneously indexing choice difficulty. Two additional studies – one by Boorman and colleagues¹⁹, the other a more recent study from Kolling and colleagues²⁰ – employed tasks that shared properties with KBMR’s. In both studies, participants made risky choices partly based on explicit values manipulated by the experimenter and partly based on an inherent bias toward the safest (most probable) of the outcomes available. Again their key analyses show that dACC activity increased with the value of the non-default option (or conversely decreased with the value of the default). We simulated choices in both of these contexts, including approximations to all relevant decision-making parameters (including bias), and found that choice difficulty increased under the relevant conditions in which these experiments found increasing dACC activity (Methods 6; Supplementary Figs. 6 and 7).

Taken together with our neuroimaging findings, these analyses make clear the importance of properly estimating choice difficulty in studies of reward-based choice. Of course, difficulty is not the only domain-general decision parameter that needs to be accounted for when attempting to relate dACC activity to valuation in the context of decision-making. We describe additional factors that may have contributed to the foraging value findings in KBMR and some subsequent work in Supplementary Figs. 7–9. For instance, in KBMR’s data we found foraging value to be correlated with the degree of surprise participants might have experienced when encountering their option set (based on past experience), and found that this surprise signal correlated with activity in a rostral region of dACC in both of our experiments (Methods 7; Supplementary Fig. 9), consistent with previous context-general accounts^{2, 3, 8, 32–35}. We also highlight potential concerns arising from ROI selection³⁶, as well as some fundamental inconsistencies in the predictions made by foraging accounts across the relevant papers (Supplementary Figs. 7–8). All in all, when considered alongside our findings on decision difficulty, these additional considerations only further undermine the theoretical conclusions of the papers in question.

Discussion

KBMR's foraging account of dACC predicts that the region tracks the value of the current non-default option (e.g., switching to a new patch or performing a different task), irrespective of and potentially obviating any role the dACC has in tracking choice difficulty. We have provided strong evidence that refutes this prediction, showing instead that dACC responds to foraging value only to the extent that it offers a proxy for choice difficulty. This conclusion resonates with two closely related theories of dACC function, which focus respectively on conflict monitoring and value comparison. Conflict monitoring accounts^{14, 15} predict that dACC should track one's level of indifference in a decision-making task because higher-indifference trials require the allocation of greater control, and may be aversive for the same reason. The value comparator account⁴ similarly predicts greater dACC involvement on more difficult choice trials, but instead because this region is assumed to be directly involved in the process of comparing the values of one's options (using an accumulator model similar to the one used to model the present data). We do not purport to adjudicate between these two accounts with the current data, and note that doing so may in general be difficult, as the two make a number of overlapping predictions (see Refs. 4, 14, 37).

Both conflict and value comparator theories share in common the assumption that dACC's role in decision-making and/or cognitive control is general rather than specific to a particular decision-making context. One implication of our findings is therefore to argue against a specific role for dACC in foraging-like decisions. Rather, to the extent that dACC is responsive to a "non-default" option, our results are consistent with previous theories which propose that this reflects its role in engaging the control processes needed to *override* the (typically more automatic) default option. We recently described an integrative theory of dACC function, which proposed that the dACC is responsible for estimating the expected value of control-demanding behaviors and selecting which to execute (EVC theory²). Like KBMR, this theory predicts that dACC activity should track the expected reward for engaging in non-default behavior, inasmuch as these can be considered to be control-demanding. However, the EVC theory specifies that determining the expected value of a controlled behavior also requires estimating the demands for control as well as the costs of control itself. Both of these quantities should correlate with choice difficulty, whereas the value of the outcome (i.e., the expected reward) can in principle be dissociated from difficulty (as in Exp 2). As we have shown, KBMR's findings provide evidence only for dACC's role in encoding the costly or demanding nature of control (or the extent of comparison necessary for difficult choices), and not its value per se.

The present work also illustrates how a widely used quantitative model of decision making can provide an estimate of choice difficulty in the presence of strong biases toward one option (or one attribute¹⁹) — a situation that otherwise risks confounding difficulty and value. In such cases, there is guaranteed to be a range of values for which the relative value of the less favored (i.e., non-default) option appears to be positive and increasing, simultaneous with the difficulty of the decision. (As discussed in the Introduction, the extreme case of this is represented by optimal behavior in a classic patch-leaving foraging task.) The degree to which this range of options dominates in a given study can be evaluated

qualitatively by comparing relative value and RT distributions (in KBMR's study these showed a strong positive relationship; Fig. 2b). Extracting this information from RTs further requires that the subject respond immediately upon reaching a decision (non-default valuation studies, as with many other neuroimaging studies, typically forbid responses within an initial window of approximately 2–4s^{17, 19, 20}). More generally, our findings highlight the risks of estimating choice difficulty in value-based decision tasks based on objective value (e.g., points), without taking account of behavioral data (i.e., choice and RT distributions). Here as well, the use of a formal model quantitatively fit to the behavioral data can be helpful.

In summary, our study corroborates the importance of dACC in foraging decisions, but not for the reasons reported by KBMR. Rather than reflecting the value of the foraging option itself, their findings and ours suggest that dACC activity can be most parsimoniously and accurately interpreted as reflecting choice difficulty or a correlate thereof (such as total evidence accumulated), as has been observed in a large number of other contexts. Importantly, in a changing environment, indifference may mark the optimal point of transition between default and foraging behavior, and thus explain its engagement in the context of the task introduced by KBMR. These observations should help unite the large existing literature on dACC function with the more recent one that has begun to emerge concerning its involvement in foraging decisions.

Online Methods

1. Participants

Healthy right-handed individuals were recruited to participate in a neuroimaging study involving choices and rewards. 15 individuals completed Experiment 1 (9 female; $M_{\text{age}} = 23.5$, $SD_{\text{age}} = 4.1$) and 14 independent individuals completed Experiment 2 (8 female; $M_{\text{age}} = 20.6$, $SD_{\text{age}} = 2.4$). Additional participants were excluded a priori for excessive head movement (1), incomplete sessions (4), misunderstanding instructions (assessed during a structured post-experiment interview; 2), or not meeting KBMR's criterion of foraging on at least eight trials (6). Of the 14 included participants in Exp 2, we excluded one of three trial blocks for one participant who reported falling asleep during that block. No statistical tests were used to pre-determine sample sizes but our sample sizes are within the standard range in the field. Participants provided informed consent in accordance with policies of the Princeton University institutional review board.

2. Procedure

Experiment 1 followed the procedure described by KBMR¹⁷ (Fig. 2a). Briefly, participants learned fixed reward values associated with 12 abstract symbols. Participants then performed a decision task that proceeded in two stages. In Stage 1, the participant chose whether to *engage* a pair of symbols offered or to *forage* for a better pair from a set of six other symbols also shown. Each time they chose to forage, the current engage pair was swapped out for a random pair from the forage set. Foraging came with an explicit *search cost* that was indicated at the start of Stage 2 by a colored box surrounding the search set, and feedback regarding this cost was provided within the delay period that was also incurred

when choosing to forage. Search costs were associated with point losses of varying (but known) magnitudes, and were incurred on 70% of forage choices; the remaining forage choices were associated with no point loss. Once the participant chose to engage the offered pair (and no longer forage, or not to forage in the first place), they proceeded to Stage 2. In Stage 2, each of the two symbols was independently and randomly assigned an explicit probability of success (range: 20–90%) and the participant chose which of the two symbol-probability pairs (gambles) they would like to play. If the gamble chosen was successful, the participant received the full reward value associated with that symbol. Participants were given feedback both after each forage choice (regarding whether they incurred the search cost) and after making their Stage 2 choice (regarding the success of both the chosen and unchosen options). Rewards received were tallied by a progress bar shown at the bottom of the screen; each time this bar filled to a goal line, the participant received \$1.00 and the bar restarted.

Jittered Poisson-distributed intertrial intervals (ITIs) were added between choosing to forage and receiving feedback regarding whether the search cost was incurred on that trial (range: 2–6s; $M = 3.0s$); between foraging feedback (1–2s duration) and the next Stage 1 choice set for that trial (range: 2–4.5s; $M = 2.7s$); between choosing to engage and being shown the engage probabilities (range: 3–8s; $M = 4.5s$); between making a Stage 2 choice and being given feedback (range: 3–8s; $M = 4.5s$); and between Stage 2 feedback and beginning Stage 1 of the next trial (range: 2–4.5s; $M = 2.7s$).

Relative to KBMR's experiment, our procedure only differed in that participants in our study were able to submit a response as soon as they were provided with the relevant information at a given stage. KBMR, by contrast, included a jittered "monitor phase" at the start of each stage that prevented participants from responding for 2–4s at the start of Stage 1 and 1–4s at the start of Stage 2, resulting in truncated RT distributions that preclude modeling with the drift diffusion model (DDM; see section 4.2). In order to maintain similar overall timing, we buffered ITIs immediately following Stage 1 and Stage 2 responses by an additional 2.7s or 1.8s (the average length of the respective monitor phases) minus the RT on that trial (i.e., any extra time that would have been captured by the monitor phase during the response period was instead added to the ITI, without the participants' knowledge).

As in the original experiment, Exp 1 included 135 trials, broken up into 2 blocks, and the 6 forage and 2 engage values presented at the start of each trial were drawn at random, without replacement, from a uniform distribution of the 12 total symbol values. This necessarily resulted in an over-representation of trials in which the average forage and engage values were similar, relative to those in which forage values were much higher or much lower than engage values. In order to explore the distribution of relative forage values more fully, particularly those trials in which the relative value of foraging exceeded the individual's subjective indifference point, Exp 2 made a number of modifications.

First, we used a much wider distribution of potential reward values (69 rather than 12). In order to accomplish this without imposing excess memory load on our participants, we used explicit numeric reward values (2–70) during the task rather than abstract symbols, and all explicit reward values were scaled down by a factor of 10 relative to Exp 1 (where the range

shown was 20–130 points). These scaled-down numeric values mapped onto similar monetary reward values as in Exp 1 (e.g, the point value ‘100’ in Exp 1 was associated with similar monetary reward as the point value ‘10’ in Exp 2) – we therefore apply this constant scaling factor to Exp 1 values during modeling to allow more direct comparison between the studies. Search costs were also indicated numerically rather than with the color of the search box. Furthermore, rather than uniform random sampling of forage and engage values on each trial, forage and engage values were preselected across trials so as to uniformly sample a range of relative forage values. Specifically, for each session we randomly sampled sets of forage and engage values until we found ten within each of twenty relative forage value ranges (where relative foraging value is defined as the difference between the average of the initial foraging set and the average of the pair of initial engage options) evenly spaced between approximately –29 and 56. This resulted in a total of 200 trials, performed across 3 task blocks. An asymmetric value range was used to account for the strong baseline bias to engage observed in the original task. Trial order was randomized and, aside from this initial arrangement of Stage 1 values, all other sampling of values (e.g., forage/engage options after choosing to forage, probabilities in Stage 2) proceeded as in Exp 1.

In order to slightly shorten and reduce the complexity of the individual trials, Exp 2 also used deterministic search costs rather than including a possibility that the search cost would not be incurred. As a result, the feedback portion of the forage delay was removed and the ITI between choosing to forage and receiving the next forage/engage options was set to a range of 2.3–7s ($M = 3.0$). The range of search costs used was otherwise the same as in Exp 1 (appropriately scaled, as described above). In spite of these differences, behavior was qualitatively similar in the two studies (see Supplementary Fig. 1). While average Stage 1 RTs were faster in Exp 2 (two-sample $t(27) = -2.3$, two-tailed $p = 0.030$), possibly owing to the larger number of trials and/or numeric rather than symbolic reward (reducing the need for episodic memory retrieval), the average range of RTs was similar ($t(27) = 0.71$, $p = 0.48$), as was the pattern of regression estimates for relevant decision variables for both task stages. While Exp 2 participants also exhibited a smaller overall bias, individual difference analyses performed by KBMR (see also ref. ²⁰) indicate that their account predicts that these participants should if anything encode foraging value more strongly (i.e., exhibit a stronger “readiness” to forage) than participants in Exp 1 (see legends to Supplementary Figs. 1 and 7).

Stimulus presentation and response acquisition was performed using Matlab (MathWorks) with the Psychophysics Toolbox³⁸. Participants responded with MR-compatible response keypads.

3. Image acquisition

Scanning was performed on a Siemens Allegra 3T MR system. Following KBMR, we used the following sequence parameters for volumes acquired during task performance: 3mm^3 isotropic voxels, repetition time (TR) = 3.0s, echo time (TE) = 30ms, flip angle (FA) = 87° , 43 slices, with slice orientation tilted 15° relative to the AC/PC plane. At the start of the imaging session, a high-resolution structural volume (MPRAGE) was also collected, with

the following sequence parameters: 2mm × 1mm × 1mm voxels, TR = 2.5s, TE = 4.38ms, FA = 8°.

4. Behavioral analysis

4.1. Relative value—We first estimated the relative value of the options at each stage (Stage 1: forage vs. engage, Stage 2: left vs. right). These values were used as regressors to identify regions that tracked relative value (and/or its absolute value) and, crucially, were also used as proxies for drift rate when modeling choices with the DDM (see the following section).

For Stage 2, determining relative value simply involved comparing the expected values (reward magnitude × probability) of the two options, as in the original paper. For Stage 1 we used KBMR’s primary measure of relative foraging value, which they refer to as *search evidence* (see their Eqs. S2–S5, reproduced as Eq. 1 below). Determining this quantity, involved comparing the average value of the engage pair (consisting of reward values R_1 and R_2 , whose exact probabilities P_1 and P_2 , each 55% in expectation, will be revealed in Stage 2), weighted by the ratio of the two values (in order to account for a possible preference for easier Stage 2 choices), to the average over all n possible pairs drawn from the current forage set (similarly weighted):

$$RV_{forage} = \left[\frac{1}{n} * \sum_{i=1}^n V_{offer_i} \right] - V_{Currentoffer} \quad (1)$$

where $V_{offer} = 0.55 * (w_1 * R_1 + w_2 * R_2)$, $w_1 = Pr\left(\frac{R_2}{R_1} < \frac{P_1}{P_2}\right)$, and $w_2 = 1 - w_1$

Note that RV_{forage} omits the explicit cost of foraging on any given trial. In KBMR’s study this factor was instead included as a separate regressor in their imaging analyses, so we have done the same to remain consistent with their approach. However, we note that additional analyses (not shown here) confirm that all of our findings are robust to including this cost in RV_{forage} , as estimated above. They are also robust to replacing the RV_{forage} equation above with one that makes no assumptions about offer weighting and simply subtracts the average value of engaging and the forage cost from the average value of foraging.

Following the original paper, we also performed a logistic regression to predict forage choices based on the minimum/maximum of the engage set, the search cost, and the minimum/mean/maximum of the forage set, including an intercept term. This intercept term was significantly biased toward engaging in both experiments (Supplementary Fig. 1a), consistent both with the previous results and with our finding that these explicit reward values alone fail to fully describe the inputs to the decision process. We also use this logistic regression to generate a more “model-free” estimate of choice difficulty, relative to the DDM (see Supplementary Fig. 5c). We do this by generating the log-odds of choosing to forage vs. engage on each trial based on the choice values (the ones entered into the logistic regression), and the best-fit regression weights (including the intercept). Choice difficulty was defined as the negative absolute value of this log-odds term.

4.2. Modeling behavior—In order to identify individual subjective indifference points (i.e., the points at which choices were most difficult) at each stage of the task, we fit behavioral data (choices and RTs) with the DDM²⁵. The model represents the decision process as a particle drifting towards one of two decision boundaries (e.g., forage vs. engage) with a *drift rate* that determines how much faster it moves towards one or the other (before some amount of Gaussian noise is additionally inserted into the process). The DDM, and variants thereof, have been shown to provide a reasonable approximation for value-based decision processes such as in the foraging task^{4, 26–29}, with the value of one option relative to another typically being treated as a proxy for drift rate. Notably, this model is at the heart of the *value comparator* account of dACC function⁴, which suggests that dACC activity should reflect the accumulated activity in the DDM over time. As KBMR point out when attempting to rule out both accounts, conflict and comparator theories both predict that dACC activity should be greatest when choices are most difficult. Using the DDM to identify when choices were most difficult therefore allows us to respond most directly to KBMR’s presumed lack of support for this prediction.

We used relative value at each stage (as defined above) to generate distributions of predicted choice probabilities and RTs, which were then fit to the data. We used the five-parameter version of the DDM; of these, we fixed the starting point of the decision process (0.0) but allowed the coefficient on decision noise, decision threshold, non-decision time, and value scaling factor (d below) to vary by participant and decision stage. Crucially we also included an *offset* term (c below) that was added to all relative values within a given stage when they were translated into drift rates. The equation for drift rate was therefore as follows:

$$A=(RV+c) * d \quad (2)$$

where the presumed relative value (RV) for a given choice is modified by the offset term (c ; this can also be thought of as a “latent cost”) and the scaling factor (d).

This allowed the distribution of drift rates to shift relative to the assumed distribution of relative values. In practical terms, if the assumed relative values were missing an additional fixed value (e.g. a subjective cost or bonus for foraging that was not expressed in the explicit values shown), this term would absorb that value by shifting the subjective indifference point accordingly. Like all other free DDM parameters described above, this offset term was allowed to vary by participant and trial stage.

We generated predicted choice probabilities and average RTs (threshold-crossing times) for a given set of parameters based on an analytical solution to the DDM (see appendix to ref. ²⁶). We fit the predictions of the DDM to our data in order to identify the participant and trial phase-specific DDM parameters that minimized a combination of (a) the negative log likelihood of our choice data and (b) the sum of the squared error for our RT data (after log-transforming both the actual and expected RTs). We optimized these model fits using a version of the *fminsearch* function in Matlab that implements bounded parameter ranges.

$RV_{forage-C}$ was generated by combining the original values of RV_{forage} with the best-fit offset term (c) for that individual and choice stage. Our primary measure of choice difficulty

was the negative absolute value of these corrected values of RV_{forage} ($-|RV_{\text{forage}} - C|$). Note that this measure implicitly assumes that the hidden cost estimated by our offset parameter was in a similar value currency as the explicit values that went into RV_{forage} . However, we also used the DDM to generate an alternative measure of choice difficulty that was agnostic on this front (Supplementary Fig. 5b). Similar to the approach described above for our logistic regression analysis, we used the best-fit parameters for each participant to determine the predicted likelihood of choosing one option rather than the other on each trial (i.e., likelihood of choosing to forage rather than engage in Stage 1, likelihood of choosing the option on the right rather than the left in Stage 2). We then defined choice difficulty with respect to the absolute distance between the predicted choice likelihood and indifference (i.e., 50%) for that decision, with indifference being considered the point where choices were most difficult.

5. fMRI analysis

Imaging data were analyzed in SPM8 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Functional volumes were motion corrected, normalized to a standardized (MNI) template (including resampling to 2mm isotropic voxels), spatially smoothed with a Gaussian kernel (5mm FWHM), and high-pass filtered (0.01 Hz cut-off). Separate regressors were included for the Stage 1 and Stage 2 decision phases, as well as for the feedback associated with foraging (only in Exp 1) and with Stage 2 choices. These regressors were all modeled as stick functions and the two decision phases were further modulated by parametric regressors (as described below).

We ran the following variations on this whole-brain GLM; unless otherwise noted, similar regressors were included for both stages of the task (for stage 1, we use relative value to refer to the value of forage versus engage; for stage 2, we refer to the value of the chosen versus unchosen option):

1. Parametric regressors for linear effects of relative value (RV) (Figs. 3a and 5c).
2. Parametric regressors for explicit value similarity ($-|RV_{\text{uncorrected}}|$) (Fig. 3b).
3. Parametric regressors for choice difficulty ($-|RV_{\text{corrected}}|$) (Figs. 3c and 5d; see also variants in Supplementary Fig. 5).
4. Parametric regressors for RV and choice difficulty (Supplementary Fig. 4a–b).
5. Parametric regressors for linear and quadratic factors of RV (Fig. 5c inset).
6. Parametric regressors for RT and choice difficulty.
7. Parametric regressors for RV by choice at Stage 1 (forage vs. engage; Supplementary Fig. 8).
8. Separate event regressors for each of six (Exp 1) or eight (Exp 2) binned quantiles of relative value in Stage 1 and (in a separate GLM) the same for choice difficulty. Additional bins were used for Exp 2 because of the larger number of trials and wider range of relative values tested.

For all GLMs above, Stage 2 outcomes and (when appropriate) Stage 1 outcomes were each modeled with an additional regressor. GLM #8 also included an event regressor for the onset of Stage 2 decisions, without additional parametric modulators. Also, in order to test for relative variance captured in a simultaneous regression, the default serial orthogonalization procedure in SPM was turned off for parametric regressors in GLMs #1–4, and these GLMs also included an additional regressor for the explicit forage cost on each trial (see Section 4 above). Conversely, to provide a strong test against a pure time-on-task account, GLM #6 did implement serial orthogonalization (i.e., tested for an effect of choice difficulty *after* removing variance accounted for by RT). For display purposes, all whole-brain *t*-statistic maps are shown at a voxelwise uncorrected threshold of $p < 0.01$ (corresponding to the $z > 2.3$ criterion used in KBMR's study).

We used GLM #1 and #3 for our Bayesian model comparison of choice difficulty versus foraging value in Exp 2. Since we were comparing the two accounts for dACC's role in foraging choices specifically, these GLMs were modified only to exclude parametric modulators from Stage 2. To perform these model comparisons, we first employed SPM's Bayesian equivalent of the first-level analyses described above³⁹, which produced within-participant log-evidence maps for each GLM. These log-evidence maps were aggregated into a formal Bayesian model comparison at the group level with a random-effects analysis⁴⁰ that produced voxel-wise estimates of exceedance probability for each of the models being compared. An additional Bayesian model comparison related the difficulty and foraging value models to a “baseline” model in which all of the relevant task events were modeled with indicator functions (as with all of our GLMs) but no additional parametric modulators were included.

For our sliding window analysis of foraging value (Supplementary Fig. 3), we used a variant of GLM #1 whereby only a certain subset of trials was modulated by the foraging value parameter. This subset was determined by rank-ordering trials by foraging value, and selecting those that fell within a certain percentile range window (e.g. 0th–50th percentiles). All foraging trials outside this window were modeled with a single indicator variable and no parametric modulation. We performed two sets of windowed analyses, one using windows of 70% per GLM, the other using windows of 50% per GLM. Consecutive windows were shifted by 10 percentiles, resulting in 86% and 80% overlap between consecutive windows in the respective analyses.

Our critical analyses (particularly in Exp 2) relied on a region-of-interest (ROI) approach. We extracted beta estimates from a sphere (9mm diameter) drawn around peak dACC coordinates from the relative foraging value contrast (from GLM #1 above; Fig 3A; MNI coordinates [x,y,z]: 4, 32, 42). While this results in a circular analysis for Fig. 5a (which is shown for visual comparison to the patterns of difficulty-related activity in Fig. 5b), we chose these coordinates to provide the strongest bias in our Exp 2 analyses of relative value (Fig. 5c and Supplementary Figs. 3, 4 and 8) in favor of detecting the same linear pattern of activity as in Exp 1. This ROI is used for all other dACC ROI analyses described in the paper, with the exception of the forage choice analyses shown in Supplementary Fig. 8e (right panel). For this analysis, an ROI was drawn around the peak dACC coordinates from Exp 2's sliding analysis of foraging value (Supplementary Fig. 3b) focusing on the 20th–70th

percentile window, which showed the strongest linear foraging value effect across the two windowed analyses (coordinates: 6, 28, 34). Unless otherwise indicated, statistical inferences for these ROI analyses was performed with two-tailed one-sample *t*-tests.

6. Simulating the role of choice difficulty in more recent studies of dACC and non-default value

A few studies have attempted to provide supporting evidence for dACC's role in non-default valuation since KBMR's study was published. In particular, Boorman et al. (2013)¹⁹ and Kolling et al. (2014)²⁰ engaged participants in different kinds of risky choice settings and exploited their default risk aversion in order to explore dACC responses to the increasing relative value of a non-default (riskier) option (see figure legends to Supplementary Figs. 6 and 7). We examined the possibility that choice difficulty might account for their relevant findings. To this end, we simulated sets of agents to make decisions in each of these choice contexts. These simulated decision-makers simply made choices based on a set of decision parameters (including decision noise) that were passed through a binary logistic (Kolling et al.) or trinary softmax (Boorman et al.) function, which simultaneously served to provide us with the relative probabilities of each of the choices. These choice probabilities were used to estimate choice difficulty, based either on the proximity to indifference (50%) for the binary choices, or the Shannon's entropy of the choice probabilities for the trinary choices. Depending on the relevant set of dACC results, we either show the average choice difficulty for different conditions (Supplementary Fig. 7b, left) or show the beta estimates resulting from regressing choice difficulty on a relevant decision variable (Supplementary Figs. 6, bottom, and 7b, middle and right).

Decision parameters and trial values were chosen to produce behavioral patterns approximating those reported for a given study (Supplementary Figs. 6, top, and 7a left/right; compare to Fig. 2D in ref. 19 and Figs. 1C/2A in ref. 20, respectively), and in each case included a parameter that served to bias choices toward the currently most probable¹⁹/safer²⁰ of the choice options. For simplicity, and without loss of generality, we assumed full knowledge of the outcome probabilities on each trial of Boorman et al.'s study, which were randomly selected on each trial rather than varying gradually over time. We also note that the qualitative patterns of results shown in Supplementary Figs. 6 and 7 were robust to variations in decision parameters that provided similar qualitative fits to the observed behavior in a given experiment. Collectively, these results comport with the intuition (emphasized in the main text) that choosing against a bias becomes easier as more reward is offered for the already-biased option and/or as less reward is offered for the alternative (non-default) option(s). In other words, choice difficulty increases alongside foraging/non-default value (within a certain range; see Fig. 1). Additional analyses from ref. ²⁰ are discussed in the legend to Supplementary Fig. 7.

7. Foraging value prediction errors

KBMR provide additional support for their foraging value account by showing that dACC activity increased with foraging value both when participants chose to engage *and* when they chose to forage (intuitively, though by no means necessarily, including trials that strongly favored foraging; see Fig. 2b). As shown in Supplementary Fig. 8, this result did

not replicate in our Exp 1 and in fact showed a robust effect in the opposite direction in Exp 2. It is therefore possible that KBMR's finding of foraging value signals for both types of forage choice was a byproduct of the choice difficulty effect we describe in the main text (and/or selection bias; see Supplementary Fig 8d). However, we noted that the skewed distribution of RV_{forage} values in the original experiment (Fig. 2b) may have also resulted in a greater proportion of unexpected/surprising (and perhaps therefore more salient) choice sets at the upper end of the RV_{forage} spectrum. Assuming such surprise occurred particularly frequently for trials on which participants chose to forage, it would have made it difficult to disentangle foraging value on those trials from surprise (cf. unsigned prediction error) at the infrequent configuration of values that would encourage foraging (for related arguments, see ref. ⁴¹). This fact is particularly relevant given that dACC has been previously associated with prediction errors^{8, 32–35} and interpreted in more basic terms unrelated to foraging value, including for the relevance of surprise as a signal of control demands^{2, 3}.

As a coarse measure of the degree to which a given choice set was likely to have generated such an unsigned prediction error signal, we simply took the absolute difference of each trial's RV_{forage} from the observed mean RV_{forage} up to that point. We found that this was in fact correlated with RV_{forage} in our replication study (Exp 1), conditional on their choice whether or not the participant chose to forage on a given trial. In particular, we see a strong positive correlation between RV_{forage} and $RV_{forage\ surprise}$ on the trials where subjects chose to forage ($M_\rho = 0.89$, $t(14) = 31.4$, $p = 2.2 \times 10^{-14}$). This correlation reversed on trials where subjects chose to engage ($M_\rho = -0.54$, $t(14) = -10.6$, $p = 4.5 \times 10^{-8}$). These estimates of foraging value prediction error were used to generate the whole-brain analyses shown in Supplementary Figure 9.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to Sam Feng for assistance in data analysis. This work is supported by the C.V. Starr Foundation (A.S), the National Institute of Mental Health R01MH098815-01 (M.M.B), and the John Templeton Foundation.

References

1. Rushworth MFS, Kolling N, Sallet J, Mars RB. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol.* 2012; 22:946–955. [PubMed: 22572389]
2. Shenhav A, Botvinick MM, Cohen JD. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron.* 2013; 79:217–240. [PubMed: 23889930]
3. Alexander WH, Brown JW. Medial prefrontal cortex as an action-outcome predictor. *Nat Neurosci.* 2011; 14:1338–1344. [PubMed: 21926982]
4. Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A. Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci USA.* 2011; 108:18120–18125. [PubMed: 22006321]
5. Shackman AJ, et al. The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat Rev Neurosci.* 2011; 12:154–167. [PubMed: 21331082]

6. Venkatraman V, Huettel SA. Strategic control in decision-making under uncertainty. *Eur J Neurosci.* 2012; 35:1075–1082. [PubMed: 22487037]
7. Holroyd CB, Yeung N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn Sci.* 2012; 16:121–127.
8. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nat Neurosci.* 2007; 10:1214–1221. [PubMed: 17676057]
9. Rushworth MFS, Noonan MaryAnn P, Boorman Eric D, Walton Mark E, Behrens Timothy E. Frontal cortex and reward-guided learning and decision-making. *Neuron.* 2011; 70:1054–1069. [PubMed: 21689594]
10. Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 2000; 23:475–483. [PubMed: 11006464]
11. Paus T, Koski L, Caramanos Z, Westbury C. Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport.* 1998; 9:R37–R47. [PubMed: 9674567]
12. Pochon JB, Riis J, Sanfey AG, Nystrom LE, Cohen JD. Functional imaging of decision conflict. *The Journal of Neuroscience.* 2008; 28:3468–3473. [PubMed: 18367612]
13. FitzGerald THB, Seymour B, Dolan RJ. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci.* 2009; 29:8388–8395. [PubMed: 19571129]
14. Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. *Psychol Rev.* 2001; 108:624–652. [PubMed: 11488380]
15. Botvinick MM. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci.* 2007; 7:356–366. [PubMed: 18189009]
16. Shenhav A, Buckner RL. Neural correlates of dueling affective reactions to win-win choices. *Proc Natl Acad Sci USA.* (in press).
17. Kolling N, Behrens TEJ, Mars RB, Rushworth MFS. Neural mechanisms of foraging. *Science.* 2012; 336:95–98. [PubMed: 22491854]
18. Pearson JM, Watson KK, Platt ML. Decision making: The neuroethological turn. *Neuron.* 2014; 82:950–965. [PubMed: 24908481]
19. Boorman ED, Rushworth MF, Behrens TE. Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *J Neurosci.* 2013; 33:2242–2253. [PubMed: 23392656]
20. Kolling N, Wittmann M, Rushworth MFS. Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron.* 2014; 81:1190–1202. [PubMed: 24607236]
21. Mobbs D, et al. Foraging under competition: the neural basis of input-matching in humans. *J Neurosci.* 2013; 33:9866–9872. [PubMed: 23739983]
22. Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci.* 2011; 14:933–939. [PubMed: 21642973]
23. Charnov EL. Optimal foraging, the marginal value theorem. *Theoretical population biology.* 1976; 9:129–136. [PubMed: 1273796]
24. Stephens, DW.; Krebs, JR. *Foraging Theory.* Princeton University Press; 1986.
25. Ratcliff R. A theory of memory retrieval. *Psychol Rev.* 1978; 85:59.
26. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev.* 2006; 113:700–765. [PubMed: 17014301]
27. Milosavljevic M, Malmaud J, Huth A, Koch C, Rangel A. The Drift Diffusion Model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgm Decis Mak.* 2010; 5:437–449.
28. Krajbich I, Armel C, Rangel A. Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci.* 2010; 13:1292–1298. [PubMed: 20835253]
29. Basten U, Biele G, Heekeren HR, Fiebach CJ. How the brain integrates costs and benefits during decision making. *Proc Natl Acad Sci USA.* 2010; 107:21767–21772. [PubMed: 21118983]

30. Grinband J, et al. The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. *NeuroImage*. 2011; 57:303–311. [PubMed: 21168515]
31. Weissman DH, Carp J. The congruency effect in the posterior medial frontal cortex is more consistent with time on task than with response conflict. *PLoS ONE*. 2013; 8:e62405. [PubMed: 23638070]
32. Wessel JR, Danielmeier C, Morton JB, Ullsperger M. Surprise and error: common neuronal architecture for the processing of errors and novelty. *J Neurosci*. 2012; 32:7528–7537. [PubMed: 22649231]
33. Garrison J, Erdeniz B, Done J. Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev*. 2013; 37:1297–1310. [PubMed: 23567522]
34. Hayden BY, Heilbronner SR, Pearson JM, Platt ML. Surprise signals in anterior cingulate cortex: Neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci*. 2011; 31:4178–4187. [PubMed: 21411658]
35. Bryden DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR. Attention for learning signals in anterior cingulate cortex. *J Neurosci*. 2011; 31:18266–18274. [PubMed: 22171031]
36. Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci*. 2009; 12:535–540. [PubMed: 19396166]
37. Yeung N, Botvinick MM, Cohen JD. The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychol Rev*. 2004; 111:931–959. [PubMed: 15482068]
38. Brainard DH. The Psychophysics Toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
39. Penny W, Kiebel S, Friston K. Variational Bayesian inference for fMRI time series. *NeuroImage*. 2003; 19:727–741. [PubMed: 12880802]
40. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *NeuroImage*. 2009; 46:1004–1017. [PubMed: 19306932]
41. Hare TA, O’Doherty JP, Camerer C, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci*. 2008; 28:5623–5630. [PubMed: 18509023]

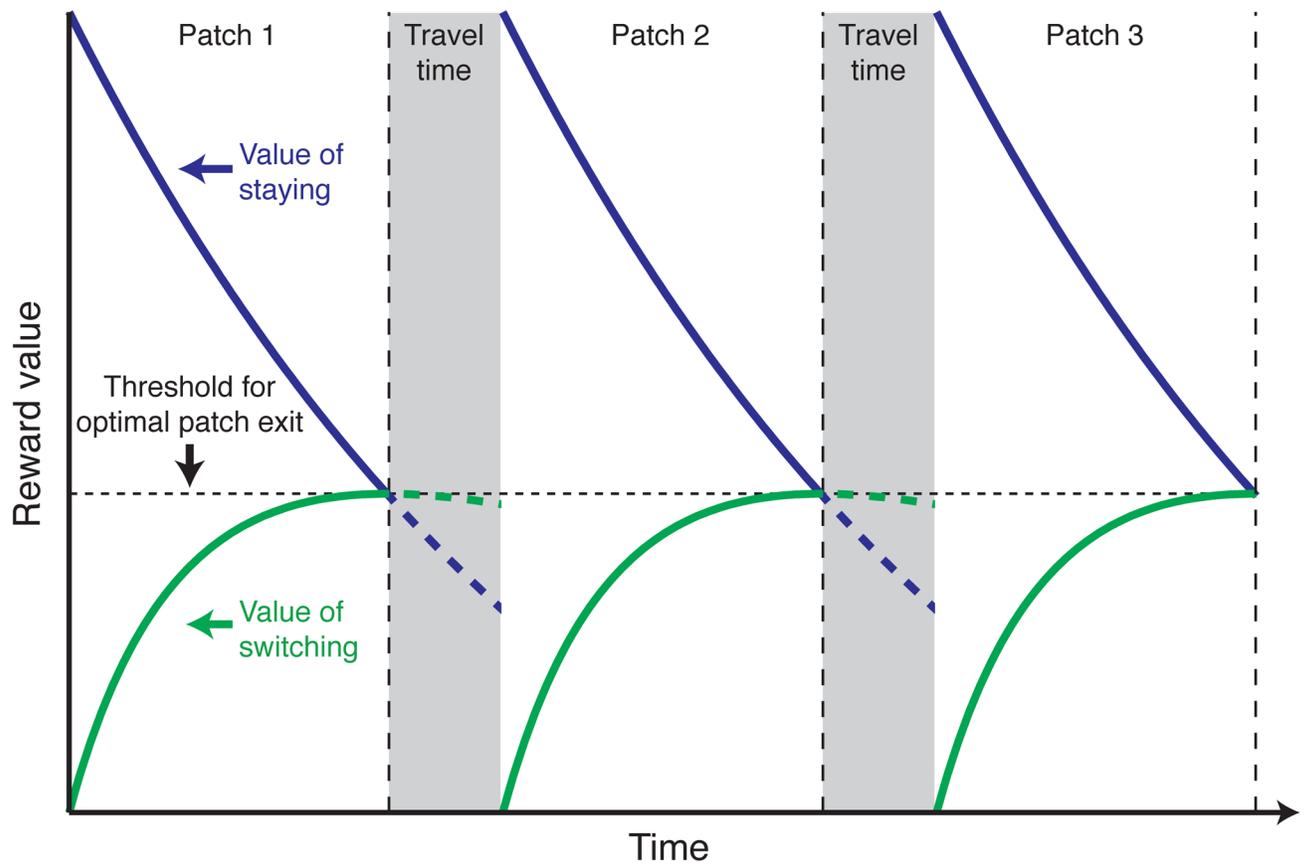


Figure 1. The role of choice difficulty in a standard foraging setting

In a typical patch-leaving scenario, an animal faces the recurring decision whether to continue harvesting a patch with decreasing marginal returns (e.g. fewer ripe fruits; blue line) or leave for a new patch. The expected value of switching to a new patch (green line) accounts for the expected reward in the new patch, as well as the travel time between patches (gray region). An optimally foraging animal will exit the patch at the point where blue and green lines meet (dashed horizontal black line), which is also their indifference point between these options. Dashed blue and green lines indicate theoretical values of staying and switching that an optimally foraging animal typically would not encounter (having already departed the patch), but that could theoretically be examined with a task like KBMR's that examines cross-sections through a foraging-like context (see Fig. 2). For simplicity, here we assume a situation where the value of each new patch, the (exponential) reward decay rate, and the travel time to a new patch remain constant across patches.

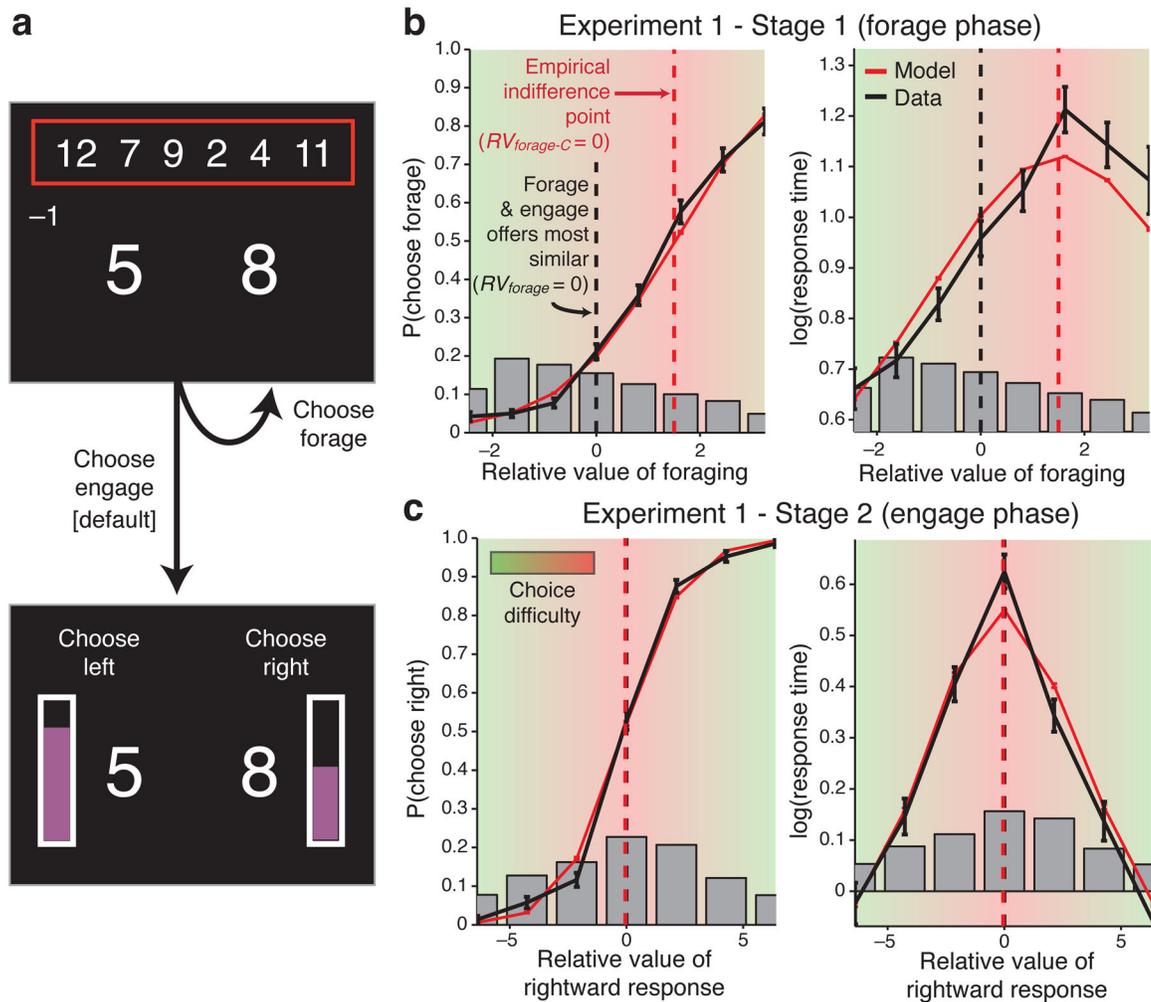


Figure 2. KBMR's estimates of value and choice difficulty do not align with behavioral data

a) Schematic of an example trial in KBMR's task. In Stage 1 (upper panels), participants are offered a pair of potential rewards (large numbers). They can choose to *forage* for a better pair of rewards from the set shown at the top of the screen (smaller numbers in the red box), in which case a random pair from that set is swapped with the current offer and they incur a forage cost (shown at left below red box) and a delay until the new choice is shown. They can forage as many times as they prefer (or not at all) before opting to proceed to Stage 2 (lower panel) and *engage* in the selected choice. At that point, a probability is randomly assigned to each reward (height of violet bar beside each number), and they choose which reward-probability pair to attempt. They receive the outcome of this gamble as points that accumulate at the bottom of the screen (not shown). While potential rewards were indicated numerically in Experiment 2 (as shown here), abstract symbols with learned reward associations were used in the original task and Experiment 1. **b–c)** Choice (left panels) and RT data (right panels) from the two stages of Experiment 1 (black curves). Gray bars show the histograms of trial frequencies. In Stage 1 (**b**), both the indifference point in the choice curve and the peak in response times exhibit a clear (and comparable) shift to the right of $RV_{forage} = 0$. In contrast, in Stage 2 (**c**), both the indifference point and RT peak coincide

with $RV_{right} = 0$. Red curves in each figure show the predicted RTs and choice probabilities, and corresponding indifference points (vertical dotted lines) based on fits of the decision making model (see Methods). We corrected RV_{forage} so that it was centered on this empirical indifference point ($RV_{forage-C} = 0$) and defined choice difficulty as value similarity with respect to this corrected measure ($-|RV_{forage-C}|$; green-red shading). These data further show that in Experiment 1, as in KBMR's study, the indifference point occurs toward the higher end of forage values tested, confounding forage value and choice difficulty. For display purposes panels **b–c** (and Fig. 4b) show the result of a fixed-effects model across all participants (error bars reflect s.e.m), but all analyses reported in the main text were based on individual participant fits. Note also that continuous data were used in all fits but are shown here binned. We have also truncated the x-axis to only show RV_{forage} bins with an average of five or more trials per participant, but show complete fits in Supplementary Fig. 2.

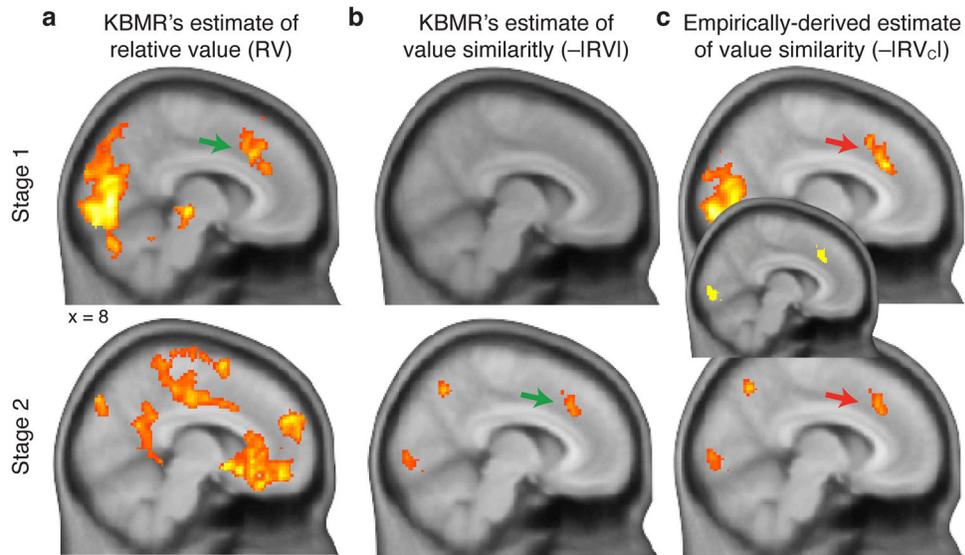


Figure 3. Experiment 1: Choice difficulty accounts for dACC activation in both stages of KBMR's original task

a) Whole-brain contrast for brain regions tracking KBMR's estimate of the relative value of forage vs. engage options in Stage 1 (RV_{forage}) (top), and the chosen vs. unchosen option in Stage 2 (bottom). **b)** Regions tracking the similarity of option values (i.e., $-|RV_{forage}|$) for the same options represented in Panel **a**. We replicate the finding of significant dACC activity in the contrasts shown in the top panel of **a** and bottom panel of **b** (indicated with green arrow), consistent with the foraging value account. **c)** However, using an estimate of value similarity corrected to align with the behavioral data (i.e., $-|RV_{forage-C}|$ in Stage 1), the same region of dACC is found to be associated with choice difficulty in both Stage 1 and Stage 2 (red arrows). A conjunction of these two contrasts (shown in the center) indicates a large degree of overlap in dACC. Statistical maps in **a–c** are thresholded at voxelwise $p < 0.01$, extent threshold of 200 voxels.

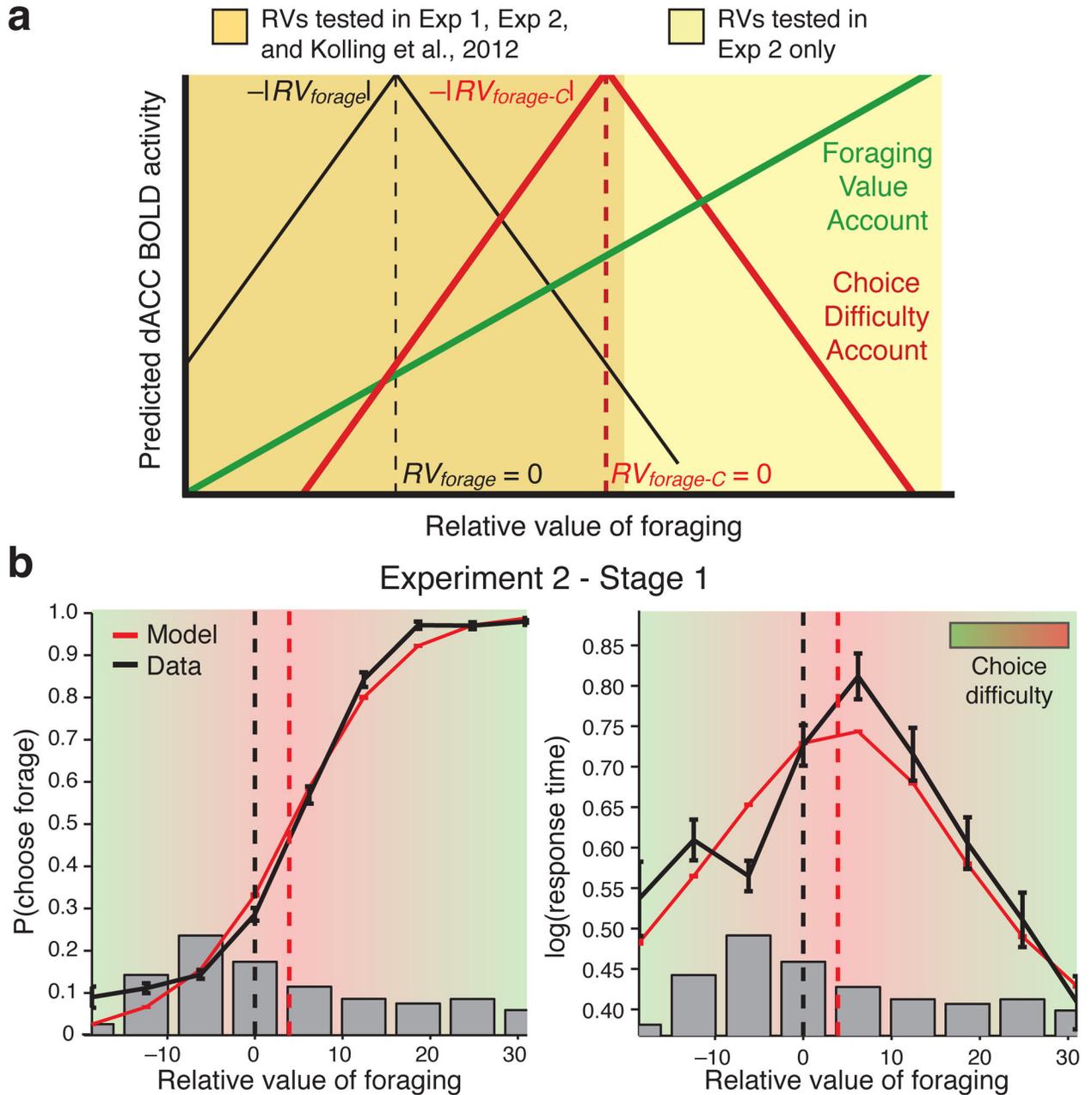


Figure 4. Experiment 2 de-confounds foraging value and choice difficulty

a) Because of the original confound (Fig. 2b), forage value (green) and choice difficulty (red) accounts make similar predictions with respect to dACC activity in Experiment 1 (Fig. 3). However, Experiment 2 dissociates the two accounts: foraging value predicts that dACC should increase monotonically, whereas choice difficulty predicts that dACC should decrease as foraging value increases beyond the indifference point (Fig. 5). **b)** Like Experiment 1, behavior and model fits from Experiment 2 exhibited a shift in the indifference point in Stage 1 (compare Fig. 2b), but a wider range of choices allowed us to de-confound difficulty and RV_{forage} .

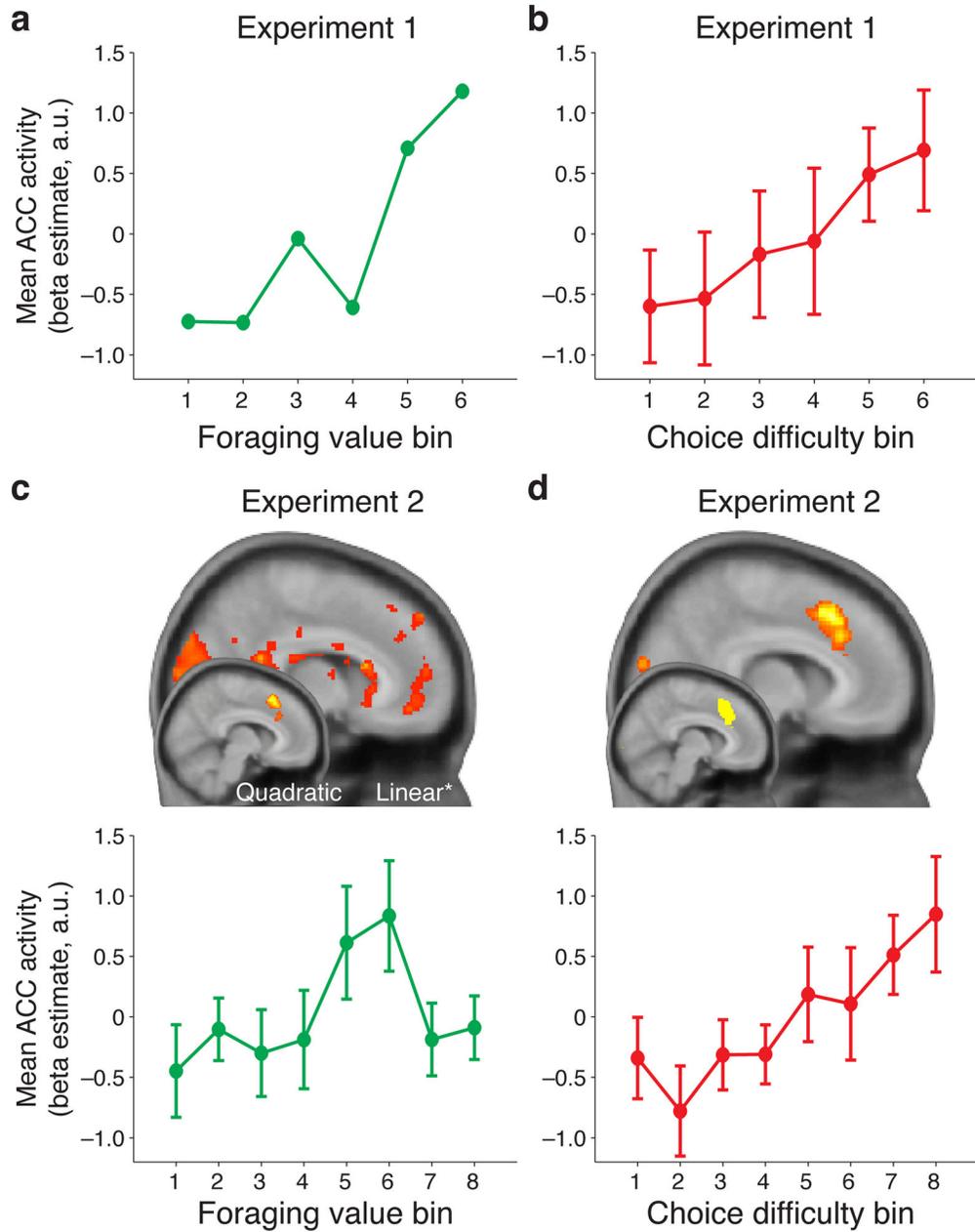


Figure 5. Experiment 2: Choice difficulty but not foraging value accounts for dACC activation when a wider range of foraging values is used

a) Average BOLD activity for each of six RV_{forage} quantiles in Experiment 1, taken from a dACC region-of-interest (ROI) around peak coordinates from the contrast shown in Fig. 3a, top panel. This is provided for visual reference but note that, unlike the remaining panels, this analysis is circular because it is intentionally biased toward the dACC region that is maximally sensitive to RV_{forage} . **b)** Given the high correlation between foraging value and choice difficulty in Experiment 1, the same pattern of activity is observed when dACC activity is binned by choice difficulty (error bars, between-subject s.e.m.). **c–d)** Results from Experiment 2, using a wider range of foraging values that orthogonalized this with respect to

choice difficulty. Images in panel **c** show whole-brain contrasts during Stage 1, showing that dACC activity exhibits a quadratic but not linear relationship to RV_{forage} ; plot in bottom panel confirms that, over the fuller range of RV_{forage} values used, dACC exhibits the non-monotonic pattern of activity predicted by the choice difficulty account (compare Fig. 4a; also see Supplementary Figs. 3 and 8). For ease of comparison with contrasts in **d**, the color map for the quadratic contrast is inverted so that negative coefficients (suggesting an inverted U-shape) appear in red-yellow. Panel **d** shows that a whole brain contrast for a linear relationship with choice difficulty again identifies dACC, plotted in bottom panel; inset shows conjunction with the same contrast for Stage 2. *This contrast is shown at a liberal voxelwise threshold of $p < 0.05$, no cluster extent threshold. All other statistical maps are shown at voxelwise $p < 0.01$, extent threshold of 200 voxels.