Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making

Romy Frömer, Carolyn K. Dean Wolf, & Amitai Shenhav

Cognitive, Linguistic, and Psychological Sciences, Carney Institute for Brain Science, Brown University, Providence, RI

Correspondence should be addressed to Romy Frömer (romy_fromer@brown.edu) or Amitai Shenhav (amitai_shenhav@brown.edu)

Abstract

How we engage with a set of options (e.g., items on a menu) is affected both by the rewards they promise and our goal in choosing between them. Typically, our goal is to maximize potential reward and minimize potential punishment. Previous research on value-based decision-making has characterized how people make decisions with this goal in mind, identifying consistent behavioral and neural signatures associated with how rewarding a set of choice options are, overall and relative to one another. However, these studies suffer from a common confound: in each case, more rewarding options were also more congruent with one's goal of choosing the best option. Previous findings can therefore be accounted for by the *reward value* or the *goal congruency* of one's options. To compare these accounts directly, we had participants make choices while their goal varied between choosing the best or the worst option, resulting in situations where either high-reward or low-reward options were most goal-congruent. We found that goal congruency uniquely accounts for past observations that decision speed varies with the overall value of one's options. Within brain regions commonly associated with choice value, fMRI activity was associated with both relative and overall goal congruency. These goal congruency correlates were dissociable from separate correlates of the overall reward associated with a choice set (independent of one's goal). Our findings call for a reinterpretation of previous research on value-based choice, and offer an approach moving forward for disentangling the roles of rewards and goals in how we evaluate our options.

Keywords: reward, decision making, emotion, functional MRI

Significance Statement

Whether it is between restaurants or career paths, to make adaptive decisions we must evaluate our options and identify those that are most conducive to our current goal. Dysfunctional decision-making can therefore result from aberrant reward processing (e.g., impulse disorders) or from aberrant goal processing (e.g., OCD, ADHD). By focusing only on how people choose their *favorite* option in a choice set (when rewards and goals are aligned), past research has been unable to distinguish the degree to which behavior and neural activity are determined by reward versus goal processing. We disentangle these processes and show that behavior and fMRI activity are differentially influenced by the promised rewards versus the degree to which those rewards align with one's current goal.

How we perceive a set of options can be affected both by the rewards we can expect from each, and by our likelihood of choosing one over another (1, 2). Previous research has demonstrated consistent behavioral and neural correlates of choice value. People make faster decisions when the value of the item they choose is large relative to the value of the item they forego (3, 4), and when the overall (i.e., average) value of their choice set is high (2, 5-7). These behavioral findings are paralleled by increasing relative and overall value related activity in regions of a well-characterized value network (8, 9), that includes regions of ventral striatum and ventromedial prefrontal cortex. However, the interpretation of these findings has been limited by a confound common to almost all such experiments, between an option's promise of reward and its ability to support a person's task goal.

Specifically, participants in studies of value-based decision-making are typically instructed to choose the item they prefer most (3, 10-14) or dislike least (14-18). As a result, the option that is most rewarding or least aversive also represented the option that is most congruent with the participant's choice goal, and the remaining options represent those that are least congruent with their goal. The relative *reward value* of a choice set (how much more rewarding the chosen option is than the alternatives) is therefore identical to the relative *goal value* of those options (how much more goal-congruent the chosen option is than the alternatives). Similarly, having more rewarding options in one's choice set (higher overall reward value) also necessarily means having more options that are congruent with one's goal (higher overall goal value). It is thus unclear whether previous behavioral findings, such as choice speeding effects of overall value, are driven by *reward* per se (5, 19-21) or by *goal congruency*. It is further unknown whether neural correlates of relative and overall value reflect their rewarding properties or their goal congruency. The present studies aim to address these open questions.

There are several a priori reasons to favor a strictly reward-based interpretation of past research on choice value. For instance, the effect of overall value on choice speed can be attributed to the Pavlovian system, which underpins our innate drive to approach rewards and avoid punishments. The Pavlovian system makes humans and other animals faster to approach rather than avoid rewards (21, 22), and makes them respond faster in environments with high rather than low overall reward (19, 20, 23). There is also ample evidence that a network of brain regions processes anticipated and experienced reward (8, 24). For instance, it has been shown that activity in this value network is correlated with the reward value assigned to items – individually (25-28) and as a set (1, 4, 9, 14) – and can be used to predict preferences between choice options (24, 29, 30). While it has been shown that preference-related activity in the value network can vary with motivational states (31) and context (32-35), it is typically assumed that this circuit tracks preferences in a *directionally-consistent* manner. That is, increased activity in this network should generally be associated with increasing reward value (36, 37). However, previous research on choice-related behavior and neural activity has failed to dissociate reward-centered accounts of these findings from accounts that are centered on the congruency of those rewards with one's current goal.

Research on recognition memory provides an excellent example of how goal congruency confounds can engender misinterpretations of a set of findings. Studies of frequency discrimination required participants to indicate which of a pair of items had been presented more frequently during an earlier study phase (38, 39). Researchers found that participants responded faster when both options were high frequency items rather than when both were low frequency items, suggesting that response speed was correlated with the overall frequency of the choice set. However, this effect reversed when participants were instructed to choose the *less* frequent item

rather than the more frequent one; when given these instructions, participants were instead faster to respond when the overall frequency was low. Thus, by manipulating choice instructions, this study dissociated item frequency from goal congruency, and showed that behavior was driven by the latter rather than the former.

Here, we employ a similar approach, asking participants to either choose the best or the worst item in a value-based choice set, to test whether task goals modulate choice value's relationship with behavior and brain activity. Across three studies, we find that a participant's decision speed depends on goal congruency, not reward value. We demonstrate that these findings can be accounted for by modifying an established computational model of choice dynamics to account for an individual's current choice goal. We further find that the brain's value network tracks relative value in terms of goal congruency rather than reward value, while at the same time also tracking overall value both in terms of goal congruency and reward value. Within striatum, correlates of (overall) reward value and (relative) goal congruency were spatially dissociated. Taken together, our findings call for a reinterpretation of previous findings in research on value-based decision-making.

Results

Participants in two studies (Ns = 30) made a series of hypothetical choices between sets of four consumer goods (Fig. 1A). Choice sets were tailored to vary in their overall value (average value of the options in the set) and in the similarity of option values, based on individual item preference ratings taken earlier in the session. Participants in Study 2 performed this choice task while undergoing fMRI.

The influence of overall value on choice speed is determined by goal congruency rather than reward value. Previous studies have consistently shown that participants are faster to select their most preferred item out of a set as the overall value of their options increases (2, 5-7). To test whether this finding is determined by the reward value of those options (as predicted by a Pavlovian account) or by their congruency with one's task goal, we explicitly varied this goal while participants made choices between sets of options. For half of the trials, we instructed participants to choose their *most* preferred from each choice set (*Choose Best* task), the standard task goal in research on value-based choice. For the other half of trials, we reversed this instruction, asking participants to instead choose their *least* preferred from a given set (*Choose Worst* task). The distribution of choice set values was held constant across the two tasks. If overall value influences choice speed through a Pavlovian mechanism, choice RTs should decrease with overall value across both tasks (Fig. 1B, top). If overall value effects are instead a reflection of goal congruency, then RTs should decrease whenever overall value is aligned with the task goal (Fig. 1B, bottom).

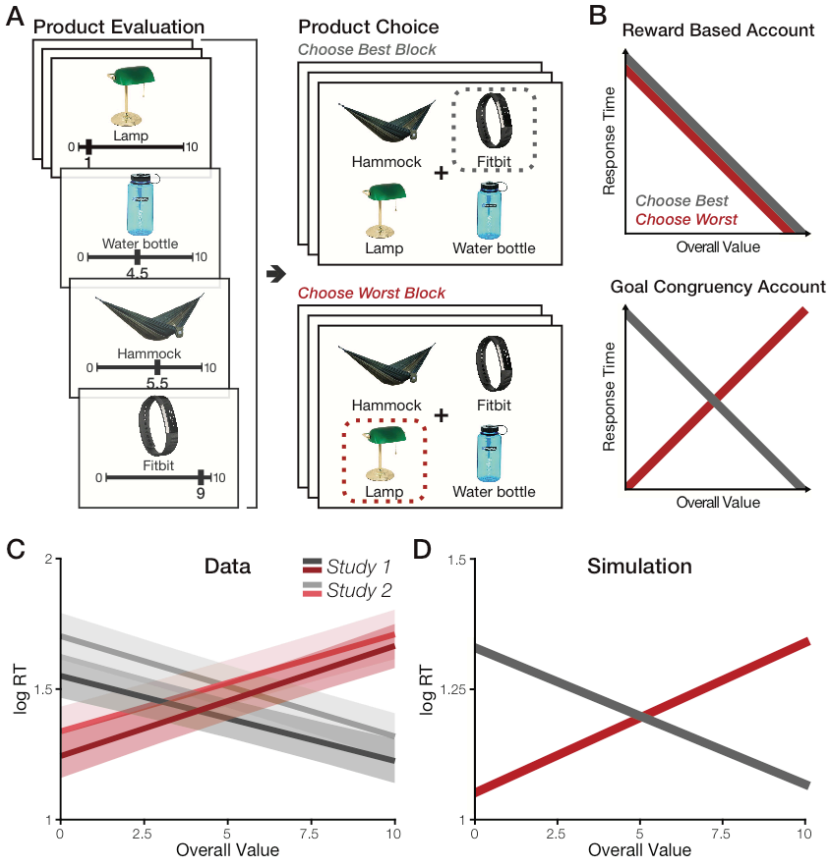

Figure 1. Overall value effects on RT are driven by goal congruency rather than reward value. A) After evaluating each item in isolation (**left**), participants saw sets of four options and (in separate blocks) were instructed to choose either the best or the worst option (**right**). The same example is shown for both blocks but each choice sets was only viewed once in a session. **B) Top:** A reward-based account predicts that RTs should decrease with overall value of the set, irrespective of the choice goal. **Bottom:** A goal congruency account predicts that RTs should decrease with overall value in Choose-Best blocks but instead increase with overall value in Choose-Worst Blocks. **C)** Both Study 1 (behavioral) and Study 2 (fMRI) find the task-specific RT reversal predicted by a goal congruency account (see also Study S1, Results 1). Shaded error bars show 95% confidence intervals. **D)** Our empirical findings were captured by an LCA model that took goal values (rather than reward values) as inputs.

Across both studies, we found strong evidence for a goal congruency account of choice behavior. Consistent with previous studies (which instructed participants to choose the best item in a set; 2, 5, 6, 7), we found that RTs decreased with overall set value for our Choose-Best task ($b <$ -0.32, $p < .001$). However, for the Choose-Worst task, we found that this relationship

flipped – RTs instead *increased* rather than decreased with overall value ($b > 0.37$, $p < .001$). Across all trials, RTs therefore demonstrated an interaction between choice goal and overall value, as predicted by a goal congruency account (Fig. 1C; Table S1, top; $b < - 0.75$, $t < -12.82$, $p < .001$). Once accounting for this interaction, we did not find residual main effects of overall value across the two tasks (Study 1: $t = 1.61$, $p = .108$; Study 2: $t = - 0.24$, $p = .809$), nor did we find a main effect of choice goal (Study 1: $t = -1.94$, $p = .062$; Study 2: $t = -0.01$, $p = .736$). In other words, depending on the task, RTs varied with the overall value either positively or negatively, but choices were not overall faster for one task or another, and overall value did not exert a stronger influence on RT in one task or another.

This pattern of findings is qualitatively consistent with the prediction that overall value influences choice RT as a function of its goal congruency rather than its association with potential rewards. In order to test this interpretation directly, we constructed two variables: one that encoded the overall reward associated with a set of options (overall reward value, OV_{reward}) and another that encoded the overall goal congruency of those options (overall goal value, OV_{goal}). OV_{reward} coded overall value as in our previous analyses, strictly increasing with the average reward value of the options. OV_{goal} coded overall value with reference to the choice goal, increasing for more rewarding sets in the Choose-Best condition and increasing for less rewarding sets in the Choose-Worst condition. We compared separate regressions predicting choice RT by OV_{reward} alone (Model 1), OV_{reward}, in interaction with task condition (Model 2), or OV_{goal} alone (Model 3). In all three models we control for main effects of task condition and VD. We found that Models 2 and 3 substantially outperformed Model 1 (Table S2). Accordingly, when including OV_{goal} and OV_{reward} in the same regression, we find a significant effect only for OV_{goal} ($b < - 0.37$, $p < .001$) and not for OV_{reward} ($-0.01 < b < 0.05$, $p > .107$; Table S1, bottom).

Importantly, this effect of overall value on choice RT is independent of the effect of value *difference*, i.e., the absolute difference between the goal value (highest item value for Choose Best and lowest item value for Choose Worst) and the remaining items. As in previous studies, we find that easier choices (those with higher value differences) are faster (Table S1), an effect that we control for in all of our analyses. Across both studies, we find that value difference also predicts the *accuracy* of a given choice – as value difference increased, participants were more likely to choose the item that achieved their current choice goal (the highest-value item in Choose Best, the lowest value item in Choose Worst; $b = 0.37$, $z > 13.14$, $p < .001$). The influence of value difference on RTs and choice accuracy did not differ between the two choice goals ($b < 0.58$, $z < 1.01$, $p > .311$). Unlike value difference, the overall value of a set did not predict choice accuracy for either task ($b < 0.49$, $z < 1.70$, $p > .089$). To demonstrate that these behavioral findings were not dependent on within-subject reference effects (between the two choice goals), we replicated our Choose Worst findings in an independent study that only included this condition (Study S1; Supplemental Results 1).

Task-dependent effects of overall value can be collectively accounted for by a goal value accumulator model. Previous research has shown that the dynamics of value-based decision-making can be captured by models of accumulating evidence to bound (3, 4, 40, 41). This work has demonstrated that a subset of such models can capture the speeding effects of both overall value and value difference on choice RTs. For instance, these effects emerge naturally from a model in which option values accumulate additively and compete with one another (the leaky competing accumulator [LCA] model (42)). However, previous implementations of this and related models of value-based decision-making have assumed that activity in this network is driven by the *reward* value of one's options, and therefore predict that overall value will only

ever result in speeding of choice RTs. Accordingly, when we simulate such a model we find that it replicates behavior on our Choose Best task (consistent with past work) but not our Choose Worst task (Fig. S1).

To account for our findings, we developed a modified LCA model. Rather than accumulating evidence of potential rewards, this model accumulates evidence in favor of one's task goal (i.e., goal value). According to this model, higher reward associated with a given option would constitute strong evidence in favor of one's goal during the Choose Best task but weak evidence in favor of one's goal in the Choose Worst task. Choice behavior simulated by this model qualitatively replicates all of our key findings across Studies 1 and 2: (1) the effect of OV_{reward} on choice RTs reverses between Choose Best trials (speeding) and Choose Worst trials (slowing) (cf. Fig. 1D); (2) OV_{goal} has a consistent speeding effect, and is sufficient to account for RTs across both tasks (cf. Fig. S2); (3) value difference has a consistent speeding effect across the two tasks (cf. Fig. S3); and (4) overall choice RTs are similar across the two tasks (cf. Fig. S3).

The brain's value network differentially tracks reward and goal value. Previous work has identified a network of brain regions – primarily comprising ventromedial prefrontal cortex (vmPFC) and ventral striatum – that tracks both the overall value of a choice set and the relative value of the chosen item (the signed difference between the values of the chosen versus unchosen items) (1, 8, 24, 28). However, because these findings are based entirely on tasks that require participants to choose the best option, it is unclear whether neural activity is being determined by the reward value of those items (as implied by past studies) or by their goal value. To disentangle these two possibilities, we had participants in Study 2 perform our Choose Best and Choose Worst tasks while undergoing fMRI. We compared the degree to which activity in a

previously defined valuation network (8) was associated with goal values and reward values, both for overall value (OV_{reward} vs. OV_{goal}) and relative value (chosen minus unchosen reward value [RV_{reward}] vs. chosen minus unchosen goal value [RV_{goal}]).

We performed separate regressions predicting BOLD activity in the value network ROI based only on reward values (RV_{reward} and OV_{reward}; GLM-1) and only on goal values (RV_{goal} and OV_{goal}; GLM-2). We found that the goal value model (GLM-2) provided a significantly better fit to neural activity ($X^2 = 5.67$, $p < 2e\text{-}16$; Table S3), with significant effects of both OV_{goal} ($b = 0.16$, $t = 2.45$, $p = .014$) and RV_{goal} ($b = 0.23$, $t = 3.50$, $p < .001$). Both goal value variables continue to be significant when including reward-related variables in the same model (GLM-3; Table S4; Fig. 2). Interestingly, though, when controlling for OV_{goal} and RV_{goal}, this full GLM revealed an *additional* effect of overall reward value ($b = 0.21$, $t = 3.17$, $p = .001$). In other words, this network tracked overall and relative value of a choice set in terms of their goal congruence, while also tracking the overall reward associated with those options (irrespective of the choice goal). Follow-up analyses showed that these results held similarly for the two constituent regions of the valuation network (ventral striatum and vmPFC), though ventral striatum was more weakly associated with OV_{goal} (Fig. S4, Table S5). None of these analyses revealed a significant correlation between BOLD activity and relative *reward* value (i.e., chosen vs. unchosen reward).

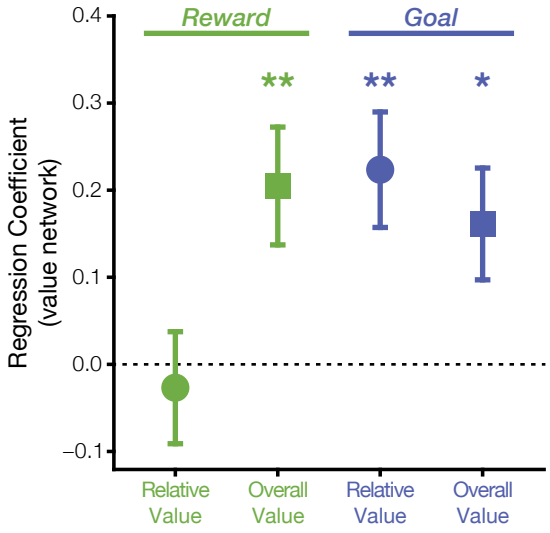

Figure 2. The valuation network tracks goal values and overall, but not relative reward value. Mixed-effects regression coefficients show that the valuation network ROI (defined a priori based on Ref. 8) tracks both overall and relative goal value, and also tracks overall (but not relative) reward value. Error bars show standard error of the mean.

Dissociable regions of striatum track reward versus goal value. Since we failed to observe relative reward value signals in our a priori ROI, we performed a follow-up whole-brain analysis (mirroring GLM-3) to test whether a subset of this network, or regions outside of it, track this variable. We did not find such correlates, even at a liberal threshold (voxelwise uncorrected $p < 0.01$, k ≥ 10 voxels). Our whole-brain analysis also revealed an unexpected functional dissociation within striatum, between more dorsal regions primarily tracking overall reward value and more ventral regions primarily tracking relative goal value (Fig. 3A). To test this dorsal-ventral dissociation directly, we defined ROIs for three sub-regions along a dorsal-ventral axis previously shown to seed distinct resting-state networks (Figure 3B; 43). We found a significant interaction between value variables and striatal sub-regions ($F(4, 25560) = 2.81$, $p = .024$), such that sensitivity to overall reward value was greatest in the dorsalmost sub-region (dorsal caudate; $b = -0.21$, $t = -2.10$, $p = .036$) whereas sensitivity to relative goal value was

greatest in the ventralmost sub-region (inferior ventral striatum; $b = 0.24$, $t = 2.19$, $p = .029$, Fig. 3C, Table S6). Sensitivity to overall goal value did not differ significantly across striatal sub-regions.

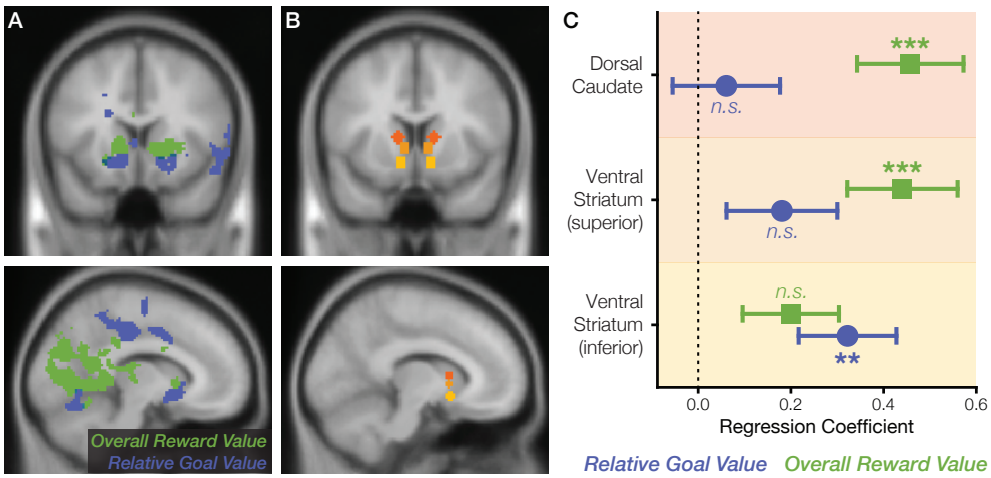

Figure 3. Reward and goal value dissociate across the striatum's dorsal-ventral axis. A) Whole brain results for relative goal value (blue) and overall reward value (green), thresholded at voxelwise $p < 0.001$ and cluster-corrected $p < 0.05$. Despite being correlated with activity in our a priori value network ROI, overall goal value did not survive the whole-brain threshold used for these follow-up exploratory analyses. **B)** To interrogate our findings across regions of striatum, we selected independent bilateral ROIs previously used as seeds for distinct resting-state networks (43): Dorsal Caudate (dark orange; [x,y,z] = +/-12, 10, 8), Ventral Striatum, superior (orange; +/-8, 10, 1), and Ventral Striatum, inferior (yellow; +/-10, 11, -9). **C)** Mixed-effects regression coefficients across these ROIs demonstrate a dorsal-ventral dissociation, with dorsal regions more sensitive to overall reward value and the ventralmost region more sensitive to relative goal value. Error bars show standard error of the mean.

Correlates of overall reward value parallel those of choice appraisal. We have previously shown that the overall value of an option set predicts how attractive a participant will appraise that set to be (*set liking*), and that these liking ratings correlate with value network activity (1, 2). Set liking correlated with activity in these regions irrespective of whether participants were instructed to appraise the set on a given trial, suggesting that these appraisals are triggered relatively automatically (1, 25, 44). Unlike previous work, our current studies were able to distinguish whether set liking was driven by overall reward value or overall goal value, and to compare the neural correlates of set liking with the neural correlate both types of overall

and relative value. Across both studies, we found that set liking correlated with overall reward value ($b > 2.83$, $t > 9.91$, $p < .001$) and not overall goal value ($b < 0.04$, $|t| < 0.74$, $p > .457$). Consistent with previous findings, in Study 2 we found that activity in the value network was correlated with set liking ($b = 0.16$, $t = 2.22$, $p = .027$). While set liking correlates did not differ significantly across regions of striatum ($F(2, 12295) = 0.72$, $p = .485$), it is notable that the qualitative pattern of responses (numerically highest in the most dorsal regions) most closely matched the pattern we observed for overall reward value (Fig. S5). Together these findings provide tentative but convergent evidence that overall reward value signals may be related to a process of automatic appraisal.

Discussion

In spite of the heterogeneity of approaches in research on value-based decision-making, a consistent set of findings has emerged with respect to the behavioral and neural correlates of overall choice set value. However, up to this point it has been unknown whether these value correlates reflect the reward value of those options (i.e., how much reward you can expect for obtaining a given option) or their goal value (i.e., to what degree does a given option support your current choice goal). To disentangle these accounts of behavioral and neural data, we had participants make value-based choices under two different choice goals, either to choose the best option or the worst option. We found that choice RTs and neural activity in value-related regions were accounted for by overall and relative *goal* value (i.e., versions of RV and OV that reverse when choosing the worst rather than the best option). We additionally identified neural (but not behavioral) signatures of overall *reward* value. Together, these results call for a significant reinterpretation of past research on value-based decision-making.

Previous studies in this field have consistently found that participants make faster decisions as the overall value of their options increases (2, 5-7, 45). However, previous work was unable to adjudicate between two accounts of these findings. One account suggests that these speeding effects are related to the rewarding properties of the choice options, because approach is overall facilitated by reward and impeded by punishment (21, 22) and/or because response vigor increases in high reward rate environments (19, 20, 23). An alternative to this reward-based account posits that speeding effects of overall value reflect increased activation of responses specific to one's goal (39, 42, 46). Our findings support the latter account, demonstrating that a choice set's overall reward value speeds choices only when the participant's goal is to choose the most rewarding item, and that it has the opposite effect when the goal is to choose the least rewarding item.

Our neural findings serve to similarly qualify reward-based accounts of neural data. Previous neuroimaging research has linked activity in a common set of brain regions to the relative and overall value of a choice set. On their face, those findings could be interpreted as reflecting a *directionally consistent* coding of potential rewards (24, 29, 36), such that (a) one's preference for a given item could be read out from activity in these regions and (b) the preferences coded in these regions could feed directly into a choice between the choice options (reflected in relative value signals). On such an account, relative value signals should always reflect whether the chosen option is generally more or less preferred than the unchosen option(s), and overall value should always reflect how much those options are generally preferred. This form of directionally consistent coding scheme is intuitive and can be seen as evolutionarily adaptive because it enables an organism to rapidly determine which options are most approach-worthy. Critically, this value code is also sufficient to perform our tasks – providing participants

just as much information about the least preferred items as it does the most preferred – obviating the need for any additional, recoded value signals. We found neural signals that are consistent with an alternate coding scheme, whereby preferences are coded relative to the current choice goal rather than being coded in a directionally consistent manner across goals. These overall and relative goal value signals would not be predicted by a reward coding account but, together with our behavioral and modeling results, are consistent with an alternate account whereby activity in these valuation networks reflects the accumulation process leading up to a choice (4, 7).

In addition to these neural signatures of goal value, we also observed neural signals that tracked the overall value of a choice set in a goal-*independent* (i.e., directionally consistent) manner. That is, the value network also tracks how much reward can be expected from the choice set, across both Choose-Best and Choose-Worst tasks. Notably, unlike our goal-dependent value signals, these goal-independent overall value signals were not accompanied by goal-independent *relative value* signals (i.e., chosen minus unchosen reward), as would be expected if this circuit was comparing items according to their reward value. It may therefore be parsimonious to interpret these overall reward signals as reflecting an appraisal process that occurs in parallel with the choice process (1, 2, cf. 47, or preceding it; 48) rather than reflecting a component of the decision-making process (1). With that said, we emphasize that these inferences pertain specifically to BOLD signals within vmPFC and ventral striatum, and that is entirely possible that other coding schemes would be found at the single-unit level within these or other regions.

There is another potential explanation for the overall reward value and relative goal value signals we found juxtaposed in the value network, which focuses on the *affective* qualities these variables share (1). Specifically, we and others have shown that overall reward value predicts how positively a person will feel about a set of items (1, 2) and that the value network tracks

such affective appraisals whether or not they are task-relevant (1, 25). Similarly, relative goal value is a proxy for confidence in one's decision, as well as for the ease or fluency of a task. Both of these engender positive affect (49-52) and both have been linked with activity in the value network (1, 53-57). This shared association with affect offers a salient (but non mutually exclusive) alternative to accounts of value network findings that focus on decision-making and other cognitive processes. However, further research on this circuit is necessary to tease apart affective and cognitive contributions to neural activity.

Within striatum, we found a surprising anatomical dissociation, whereby dorsomedial versus ventral regions were differentially associated with overall reward value and relative goal value. This dissociation was most strongly driven by an effect of overall reward value within dorsal caudate that was absent for relative goal value (whereas both variables were associated, to different degrees, with activity in ventral striatum). While not predicted, we speculate that this pattern may reflect the differential relationship of the two variables with affective and motivational states. Specifically, overall reward value is not only associated with affective appraisal (as noted above), but can also serve to motivate approach behavior. Accordingly, previous findings have linked dorsomedial striatum with action valuation (58-60) and motivation (61-63). It is also noteworthy that in spite of their relatively close spatial proximity, these ROIs fall within separate networks with distinct patterns of resting-state functional connectivity – the ventralmost region is associated with the limbic network and the dorsalmost region is associated with the frontoparietal control network (43, 64, 65). Therefore, an alternative account of this dorsal caudate finding is that overall reward value, being a task-irrelevant variable (cf. distractor), may elicit frontoparietal activity because of the increase in control demands. These

interpretations of our striatal findings are post hoc and therefore should be taken with caution, but they do suggest avenues for further inquiry that could shed light on this intriguing finding.

Our findings have important implications for future research on reward-based decision-making within healthy and clinical populations. We show that traditional approaches to studying such decisions conflate two types of variables: those that are specific to rewards per se and those that are related to how rewards are translated into goal achievement. We prescribe a method for pulling these two types of signals apart, and show that doing so reveals distinct behavioral and neural patterns. In addition to informing basic research into these distinct processes, our work carries important implications for research on clinical populations that exhibit aberrant reward and/or motivational processing (62, 63, 66-68). For instance, to the extent these individuals exhibit impulsive behaviors linked to reward rather than goal values, we provide an approach for uniquely identifying this variable of interest and honing in on the relevant deficits.

Method

Participants

Participants were recruited from Brown University and the general community. For Study 1, 37 participants were recruited. Of these, seven were excluded, one due to previous participation in a similar experiment, three due to incomplete sessions, and three due to insufficient variance in product evaluation, precluding the generation of sufficient choice sets. Thus, 30 participants (76.7 % female, M_{age} = 20.3, SD_{age} = 2.1) were included in our analyses. For Study 2, 31 participants were recruited and one was excluded due to an incomplete session. Thus, 30 participants (56.67% female, M_{age} = 21.87, SD_{age} = 4.40) were included in our analyses.

Participants in all studies gave informed consent in accordance with the policies of the Brown University Institutional Review Board.

Procedure

Participants first evaluated a series of products according to how much they would like to have them. Subsequently, participants viewed sets of four products and were instructed either to choose the item they most preferred (Choose-Best condition; cf. 2) or to choose the item they least preferred (Choose-Worst condition). The different choice goals (Choose-best vs. Choose worst) were instructed in two separate blocks, and the order of the two blocks was counterbalanced across participants (see Fig.1 and SI Methods for details). The general procedure in Study 2 was identical to that in Study 1, except that the choice phase was performed in the scanner.

fMRI Data Acquisition and Analysis

fMRI data were acquired on a Siemens 3T PRISMA scanner, and analyzed using a combination of SPM12 and R (see SI Methods). After standard preprocessing, BOLD estimates for each trial were extracted from our ROIs and analyzed using linear mixed effects models. Our primary fMRI analyses focused on an a priori *value network* ROI, based on a meta-analysis of the value-based decision-making literature by Bartra and colleagues (8). Secondary analyses examined three additional ROIs within striatum, based on an analysis of resting-state networks by Choi and colleagues (43).

Acknowledgments

This work was funded by a Center of Biomedical Research Excellence grant P20GM103645 from the National Institute of General Medical Sciences. The authors are grateful to Ayenna Cagaanan, Michelle Basta, and Hattie Xu for assistance in data collection, and to Scott Guerin, Uma Karmarkar, Matt Nassar, and Avinash Vaidya for helpful comments.

References

1. Shenhav A & Karmarkar UR (2019) Dissociable components of the reward circuit are involved in appraisal versus choice. *Sci Rep* 9(1):1958.
2. Shenhav A & Buckner RL (2014) Neural correlates of dueling affective reactions to win–win choices. *Proceedings of the National Academy of Sciences* 111(30):10978-10983.
3. Pisauro MA, Fouragnan E, Retzler C, & Philiastides MG (2017) Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nat Commun* 8:15808.
4. Hunt LT*, et al.* (2012) Mechanisms underlying cortical activity during value-guided choice. *Nature neuroscience* 15(3):470-476, S471-473.
5. Pirrone A, Azab H, Hayden BY, Stafford T, & Marshall JAR (2017) Evidence for the Speed–Value Trade-Off: Human and Monkey Decision Making Is Magnitude Sensitive. *Decision*.
6. Teodorescu AR, Moran R, & Usher M (2016) Absolutely relative or relatively absolute: violations of value invariance in human decision making. *Psychonomic bulletin & review* 23(1):22-38.

7. Hunt LT & Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. *Nature reviews. Neuroscience* 18(3):172-182.

8. Bartra O, McGuire JT, & Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* 76:412-427.

9. Jocham G*, et al.* (2014) Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. *Neuroimage* 100:498-506.

10. Krajbich I, Armel C, & Rangel A (2010) Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience* 13(10):1292-1298.

11. Polania R, Moisa M, Opitz A, Grueschow M, & Ruff CC (2015) The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nat Commun* 6:8090.

12. Gluth S, Rieskamp J, & Buchel C (2013) Classic EEG motor potentials track the emergence of value-based decisions. *Neuroimage* 79:394-403.

13. De Martino B, Fleming SM, Garrett N, & Dolan RJ (2013) Confidence in value-based choice. *Nature neuroscience* 16(1):105-110.

14. Blair K*, et al.* (2006) Choosing the lesser of two evils, the better of two goods: specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26(44):11379-11386.

15. Armel KC, Beaumel A, & Rangel A (2008) Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making* 3(5):396-403.

16. Frank MJ, Seeberger LC, & O'Reilly R C (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306(5703):1940-1943.

17. Plassmann H, O'Doherty JP, & Rangel A (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30(32):10799-10808.

18. Litt A, Plassmann H, Shiv B, & Rangel A (2011) Dissociating valuation and saliency signals during decision-making. *Cerebral cortex* 21(1):95-102.

19. Niv Y, Daw ND, Joel D, & Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191(3):507-520.

20. Niv Y, Daw N, & Dayan P (2006) How fast to work: Response vigor, motivation and tonic dopamine. *Advances in neural information processing systems*, eds Weiss Y, Scholkopf B, & Platt J (MIT Press), Vol 18, pp 1019 - 1026.

21. Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, & Frank MJ (2013) Frontal Theta Overrides Pavlovian Learning Biases. *The Journal of Neuroscience* 33(19):8541-8548.

22. Guitart-Masip M*, et al.* (2012) Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* 62(1):154-166.

23. Niv Y, Joel D, & Dayan P (2006) A normative perspective on motivation. *Trends Cogn Sci* 10(8):375-381.

24. Levy DJ & Glimcher PW (2012) The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* 22(6):1027-1038.

25. Lebreton M, Jorge S, Michel V, Thirion B, & Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64(3):431-439.

26. Plassmann H, O'Doherty J, & Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 27(37):9984-9988.

27. Knutson B, Rick S, Wimmer GE, Prelec D, & Loewenstein G (2007) Neural Predictors of Purchases. *Neuron* 53(1):147-156.

28. Peters J & Büchel C (2010) Neural representations of subjective reward value. *Behavioural Brain Research* 213(2):135-141.

29. Levy DJ & Glimcher PW (2011) Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31(41):14693-14707.

30. Tusche A, Bode S, & Haynes JD (2010) Neural responses to unattended products predict later consumer choices. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30(23):8024-8031.

31. Howard JD & Kahnt T (2017) Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 37(10):2627-2638.

32. Louie K, Khaw MW, & Glimcher PW (2013) Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences of the United States of America* 110(15):6139-6144.

33. Prevost C, Pessiglione M, Metereau E, Clery-Melin ML, & Dreher JC (2010) Separate valuation subsystems for delay and effort decision costs. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30(42):14080-14090.

34. Gan JO, Walton ME, & Phillips PE (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature neuroscience* 13(1):25-27.

35. Hare TA, O'Doherty J, Camerer CF, Schultz W, & Rangel A (2008) Dissociating the Role of the Orbitofrontal Cortex and the Striatum in the Computation of Goal Values and Prediction Errors. *The Journal of Neuroscience* 28(22):5623-5630.

36. Rangel A & Hare T (2010) Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology* 20(2):262-270.

37. Rangel A & Clithero JA (2014) The computation of stimulus values in simple choice. *Neuroeconomics (Second Edition)*, (Elsevier), pp 125-148.

38. Guerin SA & Miller MB (2011) Parietal cortex tracks the amount of information retrieved even when it is not the basis of a memory decision. *Neuroimage* 55(2):801-807.

39. Hintzman DL & Gold E (1983) A Congruity Effect in the Discrimination of Presentation Frequencies - Some Data and a Model. *Bulletin of the Psychonomic Society* 21(1):11-14.

40. Basten U, Biele G, Heekeren HR, & Fiebach CJ (2010) How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America* 107(50):21767-21772.

41. Krajbich I & Rangel A (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America* 108(33):13852-13857.

42. Usher M & McClelland JL (2001) The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* 108(3):550-592.

43. Choi EY, Yeo BT, & Buckner RL (2012) The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of neurophysiology* 108(8):2242-2263.

44. Levy I, Lazzaro SC, Rutledge RB, & Glimcher PW (2011) Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31(1):118-125.

45. Smith SM & Krajbich I (2019) Gaze Amplifies Value in Decision Making. *Psychological science* 30(1):116-128.

46. Botvinick MM & Cohen JD (2014) The computational and neural basis of cognitive control: charted territory and new frontiers. *Cogn Sci* 38(6):1249-1285.

47. Grabenhorst F & Rolls ET (2011) Value, pleasure and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences* 15(2):56-67.

48. Frömer R & Shenhav A (in prep.) Dissociable spatiotemporal signatures of appraisal and choice during value-based decisions.

49. Trent J, Lavelock C, & King LA (2013) Processing fluency, positive affect, and judgments of meaning in life. *Journal of Positive Psychology* 8(2):135-139.

50. Winkielman P & Cacioppo JT (2001) Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology* 81(6):989-1000.

51. Reber R & Schwarz N (2001) The hot fringes of consciousness. *Consciousness & Emotion* 2(2):223-231.

52. Winkielman P, Schwarz N, Fazendeiro TA, & Reber R (2003) The hedonic marking of processing fluency: Implications for evaluative judgment. *The psychology of evaluation: Affective processes in cognition and emotion.*, eds Musch J & Klauer KC (Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US), pp 189-217.

53. Guggenmos M, Wilbertz G, Hebart MN, & Sterzer P (2016) Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *Elife* 5:e13388.

54. Bang D & Fleming SM (2018) Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences*.

55. Schnyer DM*, et al.* (2004) A role for right medial prefontal cortex in accurate feeling-of-knowing judgements: evidence from patients with lesions to frontal cortex. *Neuropsychologia* 42(7):957-966.

56. Pochon JB*, et al.* (2002) The neural system that bridges reward and cognition in humans: an fMRI study. *Proceedings of the National Academy of Sciences of the United States of America* 99(8):5669-5674.

57. Shenhav A, Straccia MA, Botvinick MM, & Cohen JD (2016) Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice. *Cognitive, Affective, & Behavioral Neuroscience* 16(6):1127-1139.

58. Kim H, Sul JH, Huh N, Lee D, & Jung MW (2009) Role of striatum in updating values of chosen actions. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29(47):14701-14712.

59. Ito M & Doya K (2015) Distinct neural representation in the dorsolateral, dorsomedial, and ventral parts of the striatum during fixed- and free-choice tasks. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 35(8):3499-3514.

60. Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, & Schoenbaum G (2012) Reward prediction error signaling in posterior dorsomedial striatum is action specific. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 32(30):10296-10305.

61. Miller EM, Shankar MU, Knutson B, & McClure SM (2014) Dissociating motivation from reward in human striatal activity. *Journal of cognitive neuroscience* 26(5):1075-1084.

62. Koob GF & Volkow ND (2010) Neurocircuitry of addiction. *Neuropsychopharmacology* 35(1):217-238.

63. Dalley JW & Robbins TW (2017) Fractionating impulsivity: neuropsychiatric implications. *Nature reviews. Neuroscience* 18(3):158-171.

64. Yeo BT*, et al.* (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* 106(3):1125-1165.

65. Vincent JL, Kahn I, Snyder AZ, Raichle ME, & Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of neurophysiology* 100(6):3328-3342.

66. Mitchell MR & Potenza MN (2014) Addictions and Personality Traits: Impulsivity and Related Constructs. *Curr Behav Neurosci Rep* 1(1):1-12.

67. van den Bos W & McClure SM (2013) Towards a general model of temporal discounting. *J Exp Anal Behav* 99(1):58-73.

68. Buckholtz JW*, et al.* (2010) Dopaminergic network differences in human impulsivity. *Science* 329(5991):532.

Supplemental Information

Frömer, Dean Wolf and Shenhav

SI Methods

Procedure. The experiment was computerized using Matlab's Psychophysics Toolbox (1, 2). Participants first evaluated a series of products according to how much they would like to have each one. Items were presented individually, and participants used a mouse to indicate their evaluation of that item along a visual analog scale, anchored at 0 ("not at all") and 10 ("a great deal"). Participants were encouraged to use the entire range of the scale and reserve 0 and 10 for extreme values. Based on these subjective evaluations choice sets were created, such that half of the choices primarily varied in relative value (RV; e.g. best vs. average of the remaining items), while the other half primarily varied in overall set value (OV; average value of items in a set). Additional details regarding distribution of items and process of choice set construction can be found in (3). Participants then viewed sets of four products and were instructed either to choose the item they most preferred (Choose-Best condition; cf. 3, 4, 5) or to choose the item they least preferred (Choose-Worst condition). The different choice goals (Choose-Best vs. Choose-Worst) were performed in separate blocks, and the order of the two blocks was counterbalanced across participants. In Study 1, participants performed a total of 120 choices, 60 for each condition. They indicated their choice with a mouse click. There were no response deadline in the initial rating or choice. The general procedure in Study 2 was identical to that in Study 1 except that (1) choices were performed in the scanner; (2) they completed 144 choices (72 for each choice goal); (3) the interval between choices was varied across trials (2-7s, uniformly distributed); and (4) choices were made with a button press rather than with a mouse. Participants responded with the index and middle fingers of their left hand and right hand, using an MR-safe response

keypad. Response keys were mapped to locations on the screen (upper left, lower left, upper right and lower right). Participants practiced the response mappings prior to the choice phase by pressing the button corresponding to a given cued location until they gave the correct response on 15 consecutive trials.

Study S1. 19 participants were recruited were recruited from Brown University and the general community, out of which one had to be excluded due to technical issues. Thus, the final sample includes 18 participants (61.1% female, M_{age} = 20.1, SD_{age} = 3.6). The procedures in Study S1 were identical to Study 1, except that participants only performed the Choose Worst condition (120 trials).

Imaging Parameters. Scans were acquired on a Siemens 3T PRISMA scanner with a 64-channel phase-arrayed head coil, using the following gradient-echo planar imaging (EPI) sequence parameters: repetition time (TR) = 2500 ms; echo time (TE) = 30 ms; flip angle (FA) = 90°; 3 mm voxels; no gap between slices; field of view (FOV): 192 x 192; interleaved acquisition; 39 slices. To reduce signal dropout in regions of interest, we used a rotated slice prescription (30° relative to AC/PC). The slice prescription encompassed all ventral cortical structures but in a few participants omitted regions of dorsal posterior parietal cortex. Structural data were collected with T1-weighted multi-echo magnetization prepared rapid acquisition gradient echo image (MEMPRAGE) sequences using the following parameters: TR = 2530 ms; TE = 1.69 ms; FA = 7°; 1.0 mm isotropic voxels; FOV = 256 X 256. Head motion was restricted with a pillow and padding. Stimuli were generated using Matlab's Psychophysics Toolbox (Matlab 2013a) presented on a 24'' BOLDscreen flat-panel display device (Cambridge Research Systems) and were viewed through a mirror mounted on the head coil.

fMRI Analysis. fMRI data were processed using SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Raw volumes were realigned within participants, resampled to 2mm isotropic voxels, non-linearly transformed to align with a canonical T2 template and spatially smoothed with a 6 mm full-width at half-max (FWHM) Gaussian kernel.

Trial-wise ROI analyses. Preprocessed data were submitted to linear mixed-effects analyses using a two-step procedure. In the first step, we computed first-level general linear models (GLM) in SPM to generate BOLD signal change estimates for each trial and participant. GLMs modeled stick functions at the onset of each trial. Trials were concatenated across the two task blocks and additional regressors were included to model within-block means and linear trends. GLMs were estimated using a reweighted least squares approach (RobustWLS Toolbox; 6) to minimize the influence of outlier time-points (e.g. due to motion). The obtained estimates were transformed with the hyperbolic arcsine function (to achieve normality), and then analyzed using LMMs using lme4 (7) in R (8).

Of primary interest in the present study is activation in the valuation network, which was obtained using the conjunction mask from Bartra, McGuire and Kable (9). These results were followed up with separate masks for the vmPFC and vStr clusters within this ROI.

To test for gradients within vStr, we constructed spheres (r = 4mm) around seeds in three regions along the caudate's dorsal-ventral axis most strongly associated with the frontoparietal control network (left hemi coordinates: -12, 10, 8), the default mode network (-8, 10, 1), and the limbic network, respectively (-10, 11, -9; 10).

Whole-brain analyses. We complemented the ROI analyses with whole-brain GLMs. For these analyses, we computed first-level GLMs, modeling stick function at stimulus onsets, and

parametric regressors for 1) choice goal, 2) reward-related OV, 3) goal-related OV, 4) reward-related chosen versus unchosen value and 5) goal-related chosen versus unchosen value. Regressors were de-orthogonalized to let them compete for variance. As above, trials were concatenated across the two task blocks, additional regressors were included to model within-block means and linear trends, and GLMs were estimated using RobustWLS. Second level random effects analyses on first-level estimates were performed using SPM with voxel-wise thresholds of $p < .001$ and cluster-corrected thresholds of $p < .05$. Results were visualized using XJview (http://www.alivelearn.net/xjview).

Additional measures. The following measures were obtained in each experiment. With the exception of set appraisals, these additional evaluations are not of primary interest to the present study and were therefore not included in any analyses.

Affective Rating. Following the choices, participants sequentially viewed all choices again and for each set rated (1) how stressed/anxious that choice made them feel, (2) how confident they were in their choice, and (3) how much they liked the set as a whole (set appraisal).

Item reevaluation. At the end of the experiment, participants in Study S1and a subset of participants in Study 1 were asked to rate all items one more time with the same procedure as during subjective evaluation, but using a bipolar scale anchored at -10 ("dislike a great deal") and 10 ("like a great deal").

Individual difference measures. At the start of each session, participants completed the following personality inventories: the Behavioral Inhibition/Activation Scales (BIS/BAS), Neuroticism subscale of the NEO Five Factor Inventory, Intolerance for Uncertainty Scale, and Need for Cognition.

Analysis. Data were analyzed using lme4 package (7) for R (version 3.4.3, 8). RTs were analyzed with linear mixed effects models (LMMs) and choices were analyzed with generalized linear mixed effects models (GLMMs) with a binomial link function. We used the lmerTest package to obtain *p*-values based on Satterthwaite approximation of degrees of freedom. Choice consistency was determined based on whether participants chose the item with the highest value in Choose-Best, and conversely the lowest value in Choose-Worst. Analyses of choice consistency tested the probability of choosing the option that best satisfies the current choice goal (lowest-valued item for Choose-Best, highest-valued item for Choose-Worst); these analyses exclude trials for which multiple options equally satisfy these instructions (i.e., when multiple options share the lowest or highest value). Response times were log-transformed to reduce skew. Value difference and overall value were mean-centered and rescaled to a range from zero to one prior to analysis. Choice condition (Choose-Best vs. Choose-Worst) was included as a factor to test for main effects of choice goal and interactions between goal and choice values. We used a sliding difference contrast that estimates the difference in means between subsequent factor levels with the intercept (baseline value) being estimated as the mean across both conditions. Random effects were specified as supported by the data, according to singular value decomposition (11, 12).

Computational model. To directly compare reward-based and goal-based accounts of our findings, we simulated data using two versions of a leaky competing accumulator model (LCA; 13). In an LCA model, evidence for each potential response (in our case, each of four choice options) accumulates noisily over time based on its respective input (I_{Option}). As the evidence for each response accumulates, it competes with the evidence accumulated for each of the other responses, according to a mutual inhibition term (w). Over time, accumulated evidence decays

according to a leak (or forgetting) term (k). A decision is made once activation of one of the options reaches a decision boundary (z), and the time at which this threshold is crossed is recorded as the decision time. This decision time is combined with a fixed non-decision time (reflecting stimulus processing and response execution; t_0) to generate the response time for a given model simulation. Unlike the standard drift diffusion model, the LCA has additive properties that produce RTs that vary with the both the relative and overall value of its inputs (13).

We compared two versions of the LCA model to see which could better account for our findings. One version used reward values as inputs (as in previous studies; 14) while a second, modified form of this model used goal values as inputs. In the Choose Best condition, goal values are identical to reward values. For the Choose Worst condition goal values were computed by recoding reward values, such that lower reward values had higher goal value and vice versa. Both models otherwise used identical parameters (based on manual fitting to our behavioral data): z = 1.8, w = 0.5, k = 0.1, s (noise coefficient) = 0.5 and t_0 = 1s. The relevant qualitative predictions of the two models are robust to variation in these parameters. LCA simulations were performed using Matlab 2017b (Mathworks).

Visualization. Figures displaying projected values and confidence intervals of response times were generated using effects and ggplot2 packages based on the relevant LMMs. For viewing purposes, centered and scaled predictor variables are rescaled to their original values prior to plotting.

SI Results

1. The slowing effect of overall value in the Choose-Worst condition is not dependent on a Choose-Best reference point. In Study S1, participants only performed the Choose-Worst task. As in Studies 1-2, we found that RTs increased with increasing overall value (b = .52, t = 5.16, $p < .001$). As in these other studies, RTs were also faster for larger value differences (minimum value vs. average remaining, b = -.57, t = -11.10, $p < 2e-16$).

References

1. Brainard DH (1997) The Psychophysics Toolbox. *Spatial vision* 10(4):433-436.
2. Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial vision* 10(4):437-442.
3. Shenhav A, Dean Wolf CK, & Karmarkar UR (2018) The evil of banality: When choosing between the mundane feels like choosing between the worst. *J Exp Psychol Gen* 147(12):1892-1904.
4. Shenhav A & Buckner RL (2014) Neural correlates of dueling affective reactions to win–win choices. *Proceedings of the National Academy of Sciences* 111(30):10978-10983.
5. Shenhav A & Karmarkar UR (2019) Dissociable components of the reward circuit are involved in appraisal versus choice. *Sci Rep* 9(1):1958.
6. Diedrichsen J & Shadmehr R (2005) Detecting and adjusting for artifacts in fMRI time series data. *Neuroimage* 27(3):624-634.
7. Bates D, Maechler M, Bolker BM, & Walker SC (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*.
8. R Core Team (2014) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing).
9. Bartra O, McGuire JT, & Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* 76:412-427.

10. Choi EY, Yeo BT, & Buckner RL (2012) The organization of the human striatum estimated by intrinsic functional connectivity. *Journal of neurophysiology* 108(8):2242-2263.

11. Bates D, Kliegl R, Vasishth S, & Baayen H (2015) Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

12. Matuschek H, Kliegl R, Vasishth S, Baayen H, & Bates D (2017) Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94:305-315.

13. Usher M & McClelland JL (2001) The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* 108(3):550-592.

14. Colas JT (2017) Value-based decision making via sequential sampling with hierarchical competition and attentional modulation. *PLoS One* 12(10):e0186822.

Table S1

Comparison of overall reward vs overall goal value effects on log RT

	Study 1			Study 2		
	b	*t*	*p*	*b*	*t*	*p*
OV_{reward} by Choice Goal						
Intercept	1.42	42.65	**< 2e-16**	1.52	39.59	**< 2e-16**
$C_{b\text{-}w}$	-0.07	-1.94	.062	-0.01	-0.34	.736
VD	-0.49	-16.49	**< 2e-16**	-0.43	-9.42	**1.61e-09**
C_{worst} : OV_{reward}	0.42	10.12	**< 2e-16**	0.37	10.00	**< 2e-16**
C_{best} : OV_{reward}	-0.33	-7.78	**9.77e-15**	-0.38	-10.31	**< 2e-16**
OV_{goal} + OV_{reward}						
Intercept	1.42	42.70	**<2e-16**	1.52	39.59	**< 2e-16**
$C_{b\text{-}w}$	-0.08	-2.25	**.032**	-0.03	-0.86	.394
VD	-0.49	-16.49	**<2e-16**	-0.43	-9.42	**1.61e-09**
OV_{goal}	-0.37	-12.70	**<2e-16**	-0.38	-14.47	**< 2e-16**
OV_{reward}	0.05	1.61	.107	-0.01	-0.24	.809

Table S2

Model comparison for OV_{reward} and OV_{goal} effects on RT across studies.

	Study 1					Study 2				
Model	R^2	AIC	dAIC	X^2	*p*	R^2	AIC	dAIC	X^2	*p*
VD + $C_{b\text{-}w}$ (baseline)	0.22	4112				0.29	4396			
baseline + OV_{reward}	0.22	4112	0	2.72	.099	0.29	4398	2	0.04	.847
baseline + $C_{b\text{-}w}$: OV_{reward}	0.26	3956	-156	157.68	< 2e-16	0.33	4196	-202	203.31	< 2e-16
baseline + OV_{goal}	0.26	3957	1	2.59	.108	0.33	4194	-2	0.06	.809

Note: For each study, models are compared sequentially, and dAIC is the difference in AIC of each model to the previous model.

Table S3

Model Comparison for Reward and Goal Value – related BOLD activity in the Valuation Network ROI

Model	R^2	AIC/BIC	dAIC	X^2	*p*
$C_{b\text{-}w}$ + RT (baseline)	0.0590	11976/12040			
baseline + RV_{reward} + OV_{reward}	0.0591	11969/12045			
baseline + RV_{goal} + OV_{goal}	0.0596	11963/12040	-6	5.67	< 2e-16
baseline + RV_{goal} + OV_{goal} + OV_{reward}	0.0631	11955/12038	-8	10.20	.001
baseline + RV_{goal} + OV_{goal} + OV_{reward} + RV_{reward}	0.0632	11957/12046	2	0.20	.657

Table S4

Fixed Effects Summary GLM-3: Reward and Goal Values

	b	*t*	*p*
Intercept	0.00	0.01	.992
$C_{b\text{-}w}$	-0.01	-0.13	.894
OV_{reward}	2.08	3.17	**.001**
RV_{reward}	-0.02	-0.45	.656
OV_{goal}	1.60	2.49	**.013**
RV_{goal}	2.24	3.39	**.001**
RT	0.02	0.38	.707

Table S5

Fixed Effects Summary GLM-3: Reward and Goal Values within sub-regions of valuation network

	vmPFC			*vStr*		
	b	*t*	*p*	*b*	*t*	*p*
Intercept	-0.56	-7.92	6.83 e-09	-0.41	-5.83	2.49e-06
$C_{b\text{-}w}$	-0.06	-0.80	.430	0.08	0.92	.366
OV_{reward}	0.33	2.54	**.011**	0.28	2.82	.005
RV_{reward}	0.05	0.43	.668	-0.12	-1.28	.199
OV_{goal}	0.33	2.81	**.005**	0.08	0.77	.439
RV_{goal}	0.33	2.71	**.007**	0.30	3.01	.003
RT	0.20	2.15	.039	-0.17	-2.18	.036

Table S6

Changes in Value effects across dorsal to ventral regions in vStr

	b	*t*	*p*
Intercept	-0.40	-6.26	7.05e-07
Region_{linear}	-1.15	-6.28	3.35e-10
$\text{Region}_{quadratic}$	-0.08	-4.50	6.75e-06
OV_{reward} : Value	0.35	5.53	**3.13e-08**
RV_{goal} : Value	0.20	2.99	**.003**
Region_{linear} : OV_{reward} : Value	-0.14	-1.41	.157
$\text{Region}_{quadratic}$: OV_{reward} : Value	-0.21	-2.10	**.036**
Region_{linear}: RV_{goal} : Value	0.24	2.19	**.029**
$\text{Region}_{quadratic}$: RV_{goal} : Value	0.02	0.20	.844
RT	-0.21	-3.00	.005

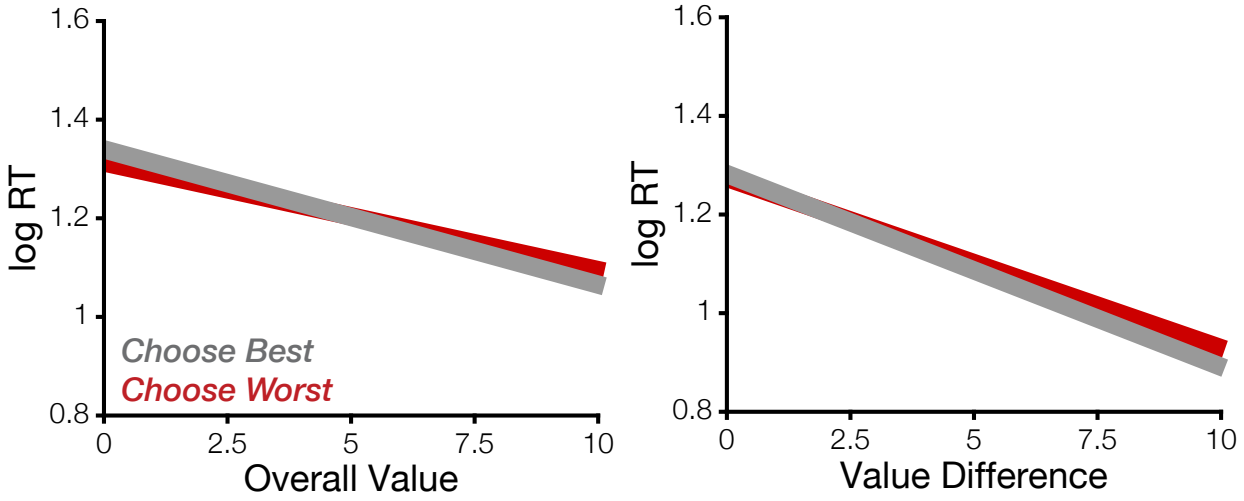

Figure S1. LCA-simulated RTs based on reward values of the items. Increasing overall value (left) and value difference (right) lead to reduced RTs irrespective of choice goal.

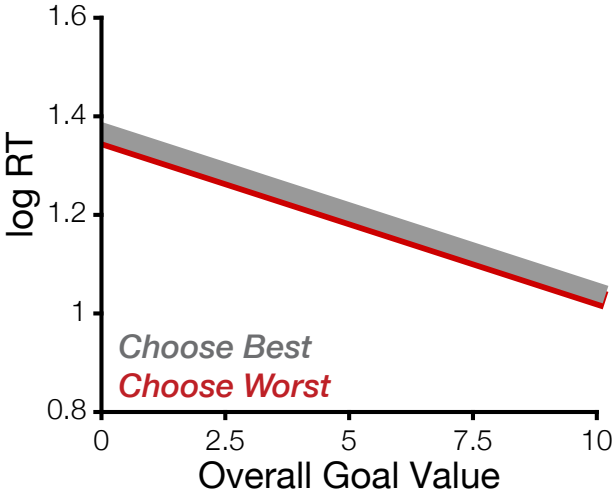

Figure S2. LCA-simulated RTs based on goal values of the items. Increasing overall goal value captures RTs speeding across both choice goals.

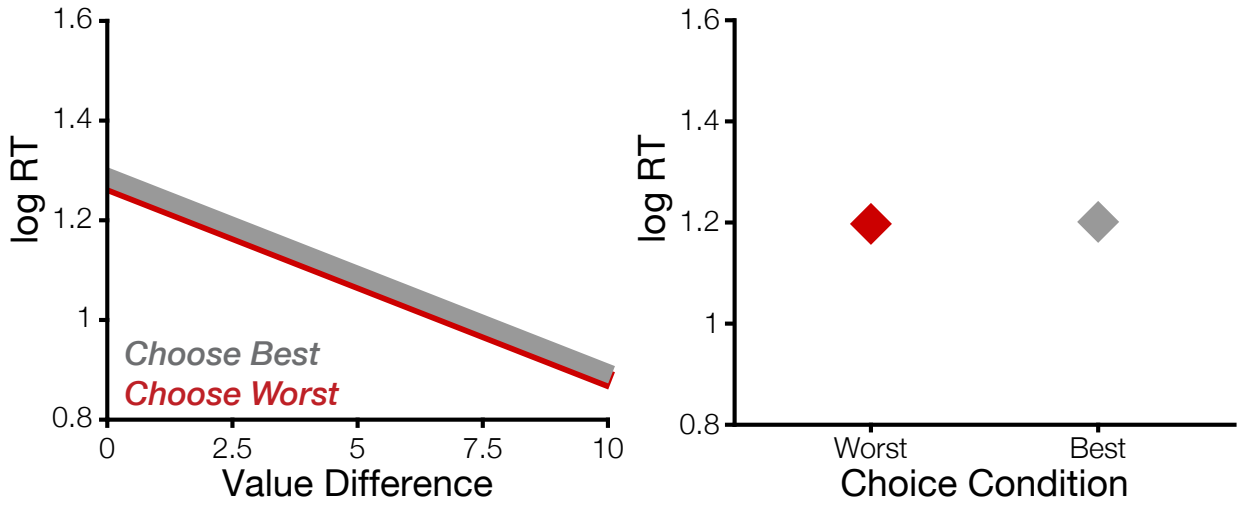

Figure S3. LCA-simulated RTs based on goal values of the items replicates patterns in observed behavior. Increasing overall value (left) leads to reduced RTs for Choose Best, but increased RTs for Choose Worst. Increasing Value Difference leads to reduced RTs irrespective of choice goal.

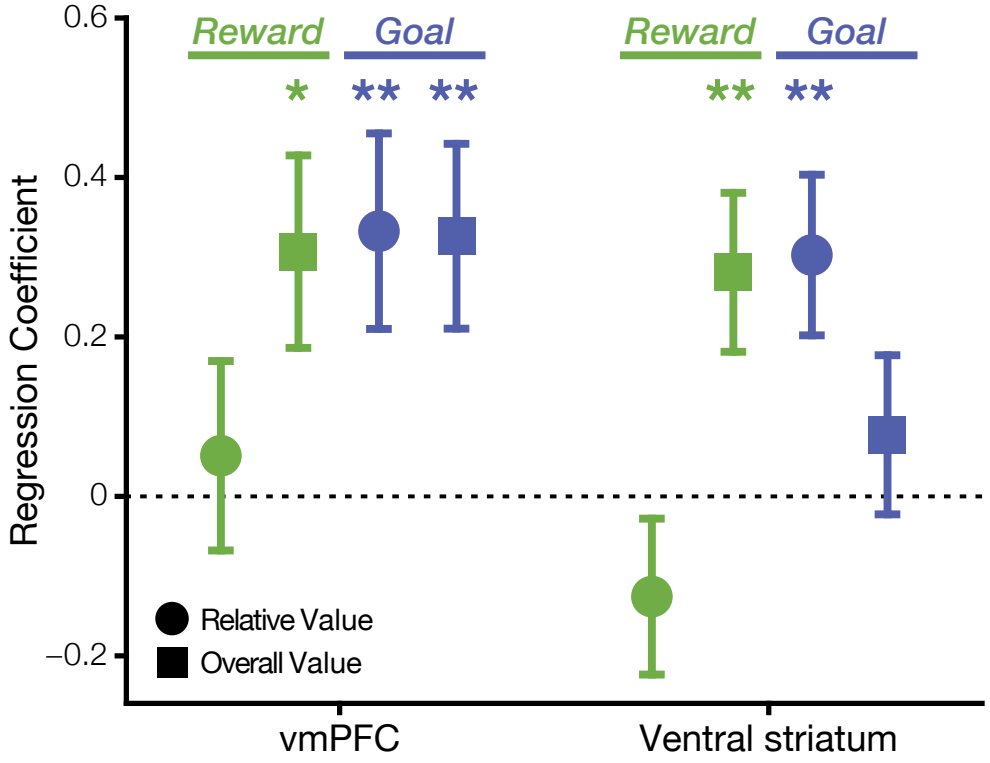

Figure S4. Effects of Relative and Overall Reward and Goal Values within constituting regions of the valuation network.

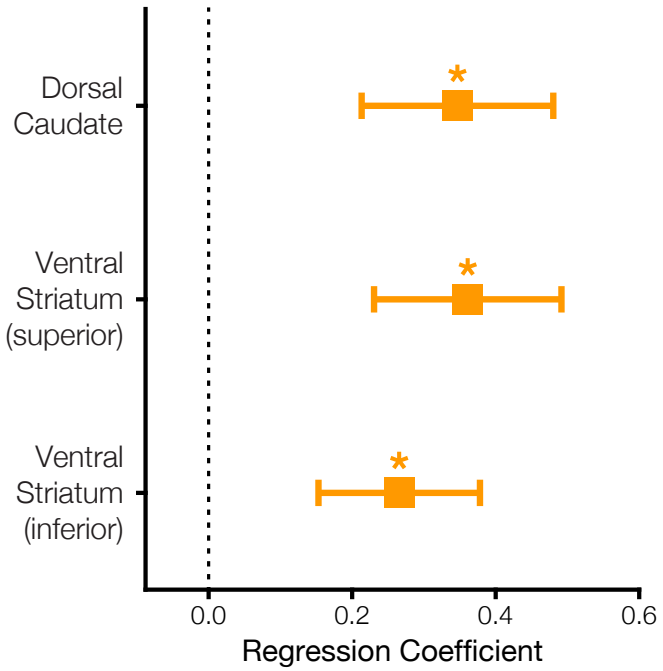

Figure S5. Set appraisal (liking) effects along the striatum's dorsal to ventral axis.