Phantom controllers: Misspecified models create the false appearance of adaptive control during value-based choice

Ritz, H.[1,2,3]*, Frömer, R. [1,2,4,5]*, & Shenhav, A. [1,2]

[1] *Cognitive, Linguistic, and Psychological Sciences, Brown University*
[2] *Carney Institute for Brain Sciences, Brown University*
[3] *Princeton Neuroscience Institute, Princeton University*
[4] *School of Psychology, University of Birmingham*
[5] *Centre for Human Brain Health, University of Birmingham*

* These authors contributed equally

Abstract

Decision scientists have grown increasingly interested in how people adaptively control their decision making. Researchers have demonstrated that parameters governing the accumulation of evidence towards a choice, such as the decision threshold, are shaped by information available prior to or in parallel with one's evaluation of an option set (e.g., recent outcomes or choice conflict). A recent account has taken a bold leap forward in this approach, suggesting that adjustments in decision parameters can be motivated by the value of the options under consideration. This motivated control account predicts that when faced with difficult choices (similarly valued options) under time pressure, people will adaptively lower their decision threshold to ensure that they make a choice in time. This account was supported by drift diffusion modeling of a deadlined choice task, demonstrating that decision thresholds decrease for difficult relative to easy choices. Here, we reanalyze the data from this experiment, and show that evidence for this novel account does not hold up to further scrutiny. Using a more systematic and comprehensive modeling approach, we show that this previously observed threshold adjustment disappears (or even reverses) under a more complete model of the data. Importantly, we further show how this and other apparent evidence for motivated control arises as an artifact of model (mis)specification, where one model's putatively controlled decision process (e.g., value-driven threshold adjustments) can mimic another model's stimulus-driven decision processes (e.g., accumulator competition or collapsing bounds). Collectively, this work reveals crucial insights and constraints in the pursuit of understanding how control guides decision-making, and when it doesn't.

Recent years have seen a growing interest in the intersections between value-based decision-making and cognitive control[1], providing researchers with fresh new insights into how people decide how to decide. These lines of research have demonstrated that decision-making is not a purely stimulus-driven process – one in which evidence accumulates based entirely on the values of one's options – but, rather, that evidence accumulation is subject to adaptive and online control. For instance, a substantial body of work has now demonstrated that people can adjust their decision threshold (how certain they must be to make a choice) by leveraging metacognitive information about this decision process. Decision-makers increase this threshold when they have conflicting evidence in support of multiple options (i.e., to buy additional time to resolve conflict and avoid mistakes[2]), and they lower (collapse) this threshold over the course of a decision, particularly when there is an impending deadline (i.e. reflecting the increasing urgency to ensure some response is made in time[3-6]). An influential new theory has proposed a bold extension on these prior accounts: that people may adjust their decision threshold not only based on metacognitive information about the task and decision process, but also based on direct access to the values of the options themselves.

Drawing on prior research on motivated control[7,8], Vassena and colleagues recently proposed that when people make deadlined choices, they make two parallel decisions based on the values of their options[9]. One of these decisions involves accumulating evidence for these options until reaching a threshold to select the best one (i.e., classical evidence accumulation). The other decision involves determining whether the values of their options merit lowering that decision threshold. To the extent that the value of one option is clearly higher than the other, their model predicts that decision-makers should maintain their typical threshold. To the extent that these option values are similar, however, their model suggests that a person should lower their decision threshold and make a more immediate decision, otherwise risking the possibility of failing to choose any option. For the same reason, a person should especially lower their thresholds ("invigorate" their response) when choosing between options of similarly high (vs. similarly low) value, to avoid an even larger loss if failing to make a choice before the deadline[10].

Vassena and colleagues' motivated control account shares properties with metacognitive accounts above, but it differs in two critical respects. First, it assumes that people are adjusting their threshold based on the specific values of their options rather than based on the environmental statistics (i.e., agnostic to the properties of each specific choice) or based on metacognitive properties of the choice (e.g., their uncertainty in their option values or in the choice process as a whole, rather than specific option values). Second, and critically, it makes the *opposite* prediction from existing conflict-based metacognitive accounts, in that it predicts that difficult choices should entail *lower* rather than higher thresholds (compare [2,11,12]).

To test this novel account, the authors had participants perform a value-based decision-making task, with each trial requiring them to select between two options within a short deadline. The authors analyzed choice behavior from this task using a drift diffusion model (DDM;[13,14]) and found that, as predicted, participants set lower decision thresholds when option values were more similar to one another (difficult choices) than when their options were more disimilar (easy choices). Additional evidence supported the further prediction that participants lowered their thresholds for high-value relative to low-value option sets[10]. Collectively, these findings offer

unique support for the proposal that people adapt their decision thresholds as a function of the expected rewards that the person is actively selecting between.

Here, we revisit this conclusion. We formally model the proposed account – that decision thresholds are adjusted based on the relative value of one's options – and show that this model poorly accounts for participants' behavior and neglects significant drivers of choice. Instead, we show that previous evidence for motivated control appeared artifactually due to multiple forms of model mimicry. First, we find that decision difficulty mimicked the influence of *overall* choice value (i.e., the average value of one's options) on choice behavior. Second, we found that joint effects across drift rate and threshold mimicked a time-varying collapse in participant's decision threshold. Finally, we found that a drift diffusion model with threshold control can mimic a control-free decision process using competing accumulators. Our findings undercut previous support for this novel motivated threshold account, and more broadly highlight significant risks for model mimicry that can occur when studying value-based decisions through an evidence accumulation framework.

Results

Vassena and colleagues (2020) had participants perform a deadlined value-based decision-making task for monetary reward (Figure 1A). On each trial, participants had a limited time window (750ms) to choose between two bundles. Each bundle contained a pair of fractals, and participants had learned the fractal values during training. For instance, the left bundle could have fractals worth 70 and 90 points, and the right bundle could have fractals worth 30 and 40 points. After they chose their preferred bundle, participants received a monetary bonus based on the point value of a random fractal from the bundle. Participants were therefore incentivized to choose the more valuable pair within the time allotted, and it was better for them to make a choice than miss the deadline (i.e., options were entirely in the domain of gains).

Vassena and colleagues argued that under these conditions it would be most beneficial for the participant to lower their response threshold when their choice is more difficult (i.e., when the bundle values are most similar to one another), because otherwise they risk failing to respond before the deadline. In other words, participants should be motivated to exert control to reduce their response threshold ("invigorate" their response) when they discover that their option values are similar to one another. In a later elaboration of their account, they unpack that this motivation should be even stronger the more valuable their options are (i.e., the higher the average value across all fractals on the screen) given that these comprise the opportunity cost of failing to respond by the deadline[10].

Models with motivated control (value-based threshold) leave room for improvement

To support their hypothesis, the authors fit a hierarchical drift diffusion-model (HDDM) to participants' accuracy and reaction times. They fit separate drift rate, threshold, and starting point parameters for easy and difficult trials, defined based on the absolute difference in bundle value (weighting the fractals equally; Figure 1B). When choices were more difficult, they found that drift rates were lower (consistent with weaker relative evidence for one option vs another[2]).

Critically, consistent with their predictions, they *also* found that participants' thresholds were lower during more difficult trials (percent of posterior density below zero: $P(x < 0) = 0.98$). Moreover, in support of the prediction that this threshold-lowering should scale with the overall value of one's options, they went on to show that thresholds were also lower for choices with high relative to low overall value[10] (controlling for difficulty).

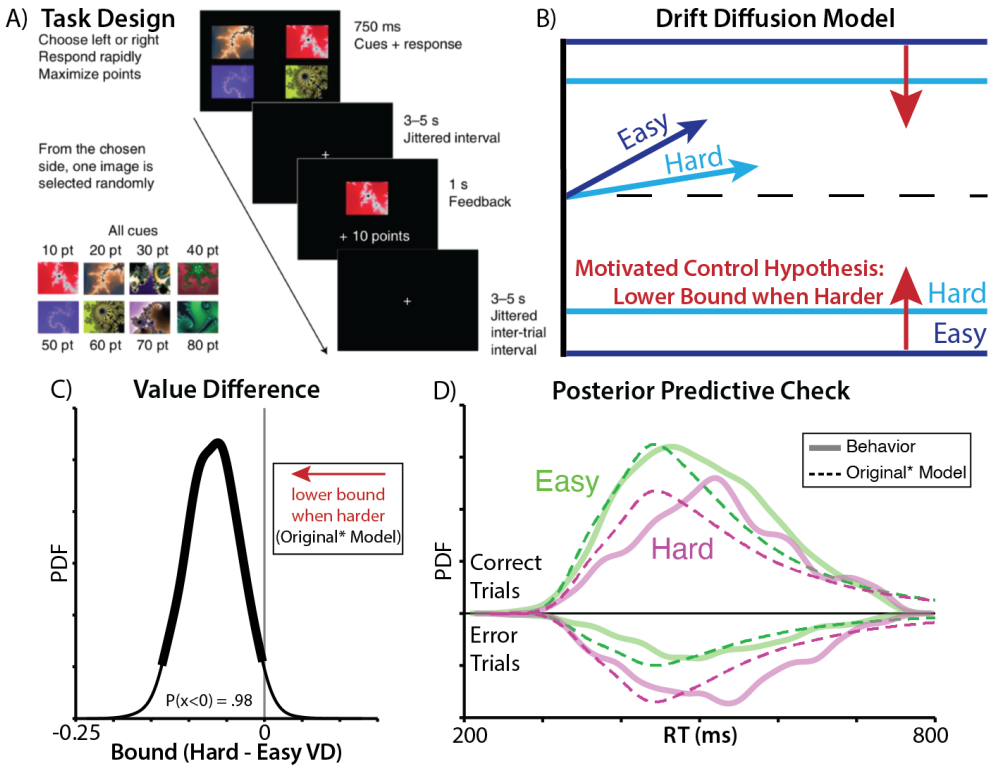


Figure 1. **Predictions of the motivated control account are confirmed when specifying the drift diffusion model as in the original study. A)** In this experiment, participants chose between two bundles of fractals worth a known amount of money. Time pressure was introduced by requiring that choices be made within a 750ms deadline, otherwise forfeiting any reward for that trial. **B)** The motivated control account predicts that harder choices (those involving similarly valued bundles; dark blue) should be associated with lower rates of evidence accumulation (as predicted by other accounts) and, critically, that these choices should also be associated with lower decision thresholds, relative to easy choices (light blue). **C)** Consistent with this account, when fitting drift diffusion models to these data using similar specifications as in the original paper, we find that harder choices are reliably associated with lower thresholds (97.5th percentile of posterior density was less than 0). **D)** When comparing RT distributions predicted by this model (dashed lines) to the actual data (solid lines), we find that the predicted distributions are substantially more skewed than the actual ones, both for easy (green) and hard (purple) choices. See also Supplementary Figure 1. Panel A adapted from [9].

Using the same HDDM model specifications provided by the original authors ('Original Model'), we replicated all the findings reported above, most notably that drift rate and threshold vary with value difference in the predicted direction (Supplementary Figure 1A). However, these analyses also suggested that there was significant room for improvement in the model's ability to capture choice behavior on this task: model-generated choices and response times, as originally specified, failed to provide a qualitative match to the empirically observed choice and RT

distributions (Supplementary Figure 1B). For instance, their best-fit model predicted RT distributions that were substantially more skewed than the actual data.

As initial steps towards addressing these concerns of model fit, we generated a revised model that encompassed the spirit of the author's original predictions, but included modest adjustments to better account for participants' choice behavior ('Original* Model'). First, to account for potential response biases (e.g., favoring the left or right option), we coded choices in terms of which side was chosen (i.e., 'response-coded' rather than 'accuracy-coded') and fit subject-specific starting points. Starting point now determined left/right biases, and not a bias towards the correct/incorrect choice on the basis on difficulty before any evidence has accumulated. Since choices were response-coded, we included the most difficult trials, which were removed in the original analysis. Second, we included an outlier term to account for attentional lapses. Third, we fit difficulty as a continuous variable, rather than a discrete variable, as the psychological construct is usually thought of as continuous and median splits (as used in the Original model) can reduce statistical power.

These model adjustments collectively led to quantitative improvements in model fit (change in DIC between Original and Original* model = 569), without altering any of the key observations: with this elaborated model, we still found that thresholds were significantly lower for difficult relative to easy choices (Figure 1C). However, when simulating distributions of choices and response times from this model, we found that these distributions still provided a poor qualitative match to empirically observed behavior (Figure 1D), predicting a much more heavily skewed RT distribution. This mismatch between model and data motivated us to consider additional revisions that would better account for core elements of the task and findings.

Evidence for motivated control is eliminated when accounting for core elements of choice

Based on the design of the choice task and the patterns of behavior that were reported by the original authors, we considered three modifications to the elaborated model above, and tested each one in turn.

First, the models above assume that participants weighted all four fractals equally when choosing between the two bundles. As a consequence, value difference (and thus choice difficulty) is determined by comparing the average value of each bundle. However, this assumption is contradicted by the authors' regression analyses showing that participants over-weight the higher-value fractal within each bundle relative to the lower-value one (placing five times as much weight in the former than the latter)[9]. We therefore tested whether separately modeling the relative and overall influence of the higher-valued and lower-valued options in each pair improved model fit.

Second, the models above only consider how parameters of the DDM will be influenced by the *difference* between two option values. However, follow-up analyses by Vassena and colleagues[10] report that these parameters also differ by *overall* set value (sum of left and right option values), consistent with a broader array of value-based decision-making research[15-18].

While the authors had originally designed their bundles to orthogonalize overall value and value difference, this was done under the assumption that participants weighed the fractals in each bundle equally. We reevaluated this assumption based on the unequal weights revealed in the analyses above, finding that these re-weighted estimates of overall value and value difference were now correlated (Supplementary Figure 2). We therefore tested whether DDM fits would be improved by including overall value as a predictor of drift and threshold in our model (along with value difference).

Third, the models above assume that participants maintained a fixed response threshold throughout the choice period. Past work suggests that decisions under time pressure are instead often better characterized by a response threshold that decreases (collapses) over time, requiring less evidence over time but guaranteeing that a response is made before one's deadline[4] (Figure 2A). A collapsing threshold model differs from the motivated control account above in that it is agnostic to the value of one's options on a given trial, but rather assumes only that thresholds collapse at the same rate on every trial (e.g., due to the deadline). We tested whether a collapsing threshold improved model fit relative to a fixed threshold model.

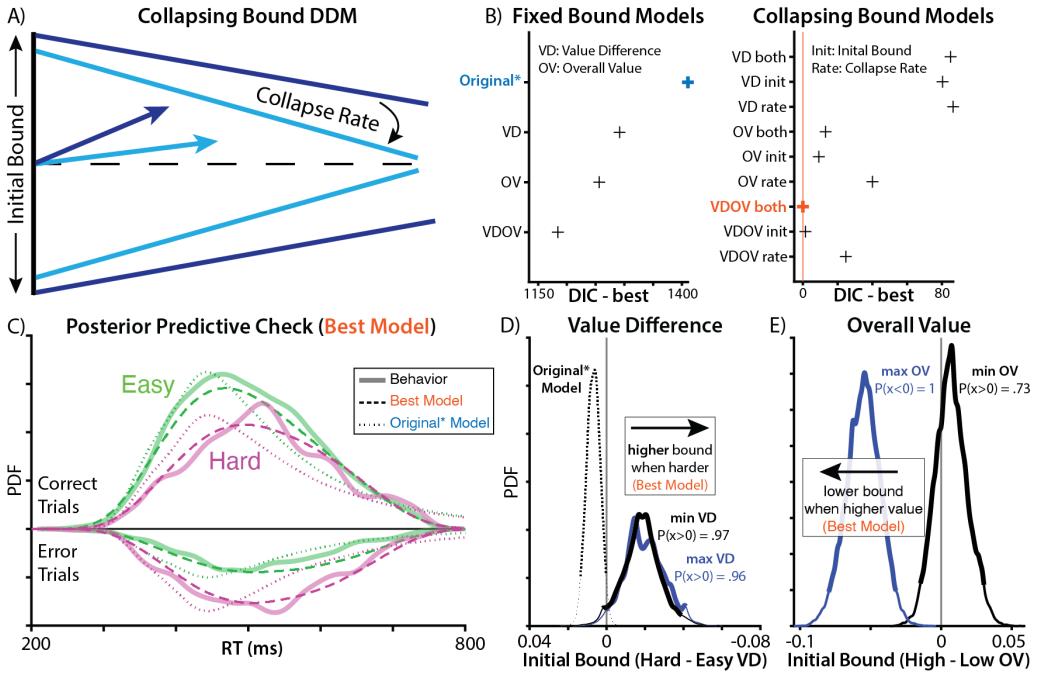


Figure 2. Elaborated models improve fit and contradict original findings. A) To account for the response deadline, we fit a DDM with a threshold that linearly decreases (collapses) over time rather than remaining fixed over an entire choice (as in the previous models). **B)** Complexity-penalized model comparisons show that incorporating this collapsing bound produces substantial improvement in model fit. Model fit further improves when allowing this collapsing bound to be further modulated by both value difference and overall (summed) value of the options, as well as when allowing for separate weights on the higher- and lower-valued fractals within each pair. **C)** Simulated behavior from the best-fitting model (dashed line) provides a substantially better qualitative fit to participant data (solid line) than the Original* model (dotted line; cf. Figure 1D). **D)** Under this best-fit model, we no longer see difficulty associated with decreases in threshold, but instead a trend in the reverse direction (compare Figure 1C). **E)** This model also shows that higher overall value is associated with lower thresholds, consistent with a separate prediction of the motivated control account.

We found that models incorporating these modifications fit participants' data far better than the Original* model (Figure 2B). A model allowing weights to differ between max and min values fit better than the original equally weighted model (ΔDIC = 119). Models with collapsing bounds substantially outperformed models with static bounds (Value difference [VD] model: ΔDIC = 1206, Overall value [OV] model: ΔDIC = 1243, model with both [VDOV]: ΔDIC = 1184). Critically, unlike models with static bounds, models with collapsing bounds were able to accurately capture the skew of participants' reaction time distributions (compare dotted and dashed lines in Figure 2C). Moreover, both in models with static bounds and collapsing bounds, models that included an influence of overall value on threshold fit better than models in which overall value was excluded (static bound VDOV vs. VD: ΔDIC = 107; collapsing bound VDOV vs. VD: ΔDIC = 85).

Whereas the Original* model had found that harder choices were associated with lower thresholds (consistent with the predictions of the motivated control account), our best-fitting models that included a value difference term now found that this variable exerted a less consistent effect on threshold (max value difference: $P(x > 0) = 0.96$, min value difference: $P(x > 0) = 0.97$). More importantly, to the extent value difference was associated with threshold adjustment, it was now in the exact opposite direction from what Vassena et al. described: harder choices were now associated with *higher* rather than lower thresholds (Figure 2D), in contradiction with the predictions of their motivated control account. Instead, the strongest influence on threshold was the overall value of the choice set (Figure 2E), more consistent with previous work on value-based decision making[15-18].

Evidence of motivated control can emerge as an artifact of model mimicry

Our findings so far undercut one major prediction of the motivated control account, which is that people use information about the similarity in their option values to adjust their threshold downward (i.e., to invigorate responses when choices are difficult). However, these models retain clear evidence for a revised account [9,19] in which people use information about the overall option value to lower their threshold (Figure 2E). This leaves open the possibility that participants are engaging in a form of motivated threshold-adjustment, but one that is more sensitive to overall value than choice difficulty. However, previous research points to an explanation of these findings without invoking any form of motivated control.

Prior work has shown similar forms of overall value-related speeding as observed in this study, showing that this behavior can naturally emerge from a stimulus-driven evidence accumulation process. In the leaky competing accumulator (LCA[20]; Figure 3A), where evidence accumulates for each response in parallel, more valuable options naturally produce stronger activations of potential responses, thus resulting in earlier threshold-crossing (see also [15,16,18]). Critically, this dynamic occurs *without* any additional control process to adjust decision thresholds. It is therefore possible that the apparent value-related control of threshold in the DDM could be mimicking an alternative decision architecture.

To test this, we simulated choice behavior in this experiment using an LCA that was coarsely tuned to mimic participants' accuracy and reaction time (See 'LCA Simulations' in Methods). Like the DDMs described above, evidence accumulation in this model was determined by the values of the individual options, with the higher value options within a bundle taking on five times the weight of the lower value options as empirically observed. The LCA makes a choice when one of its accumulators crosses a collapsing decision threshold. Critically, unlike the previous DDM-based models, value was not otherwise able to influence the decision threshold, having the same initial bound and rate of collapse across all trials.

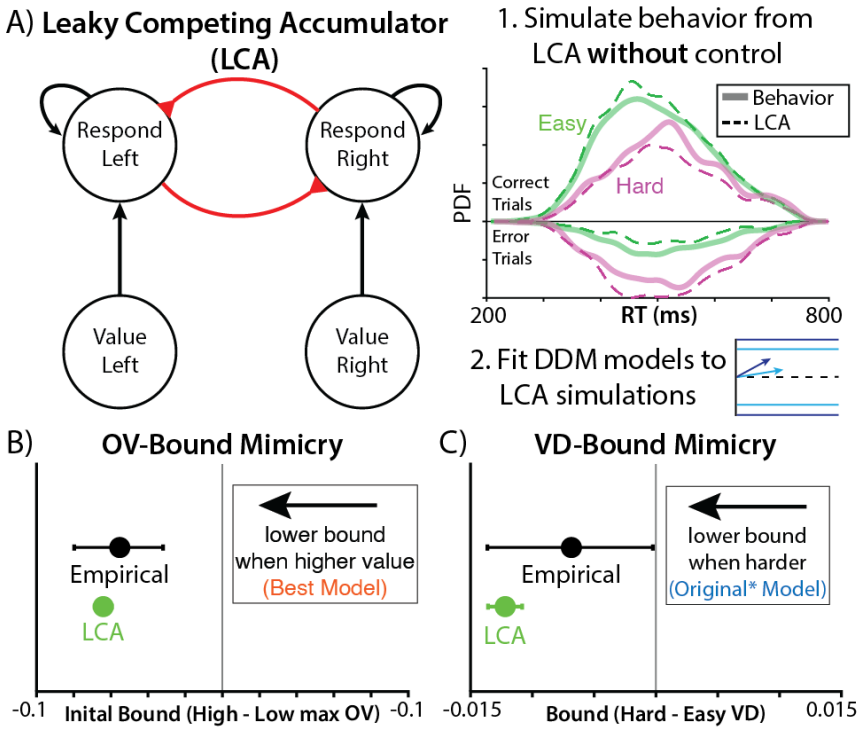


Figure 3. A control-free model can account for choice behavior while mimicking a motivated controller. A) To test whether stimulus-driven processes were sufficient to account for apparent evidence of motivated control, we simulated choice behavior from a leaky competing accumulator (LCA) that lacked a threshold control mechanism. We then fit these synthetic data to DDMs that assume that threshold is modulated by value as predicted by the motivated control account. **B)** Fitting these data to our best-fitting DDM (cf. Figure 2E), we see that the control-free LCA also produces artifactual evidence of overall-value-dependent decreases in threshold. **C)** Fitting these data to the Original* model (cf. Figure 1C), we see that our control-free LCA produces artifactual evidence of choice-difficulty-dependent decreases in threshold.

Choices simulated from our LCA exhibited all the key patterns observed in the original data, including faster and more accurate choices when there were larger value differences between the options (Figure 2A), and faster responses with higher overall set values. The fact that this "control-free" model reproduces behavior on this task indicates that stimulus-driven dynamics are sufficient to account for such behavior without invoking motivated control. It also suggests that the threshold-related adjustments that were evident in our DDM analyses may have been illusory. To test this, we took choice behavior simulated by our LCA and fit it with our DDM

models (Figure 2A). Remarkably, we once again find evidence that threshold significantly decreased with overall value, despite the *absence* of such a connection in the model that generated those data (Figure 3B). Thus, overall value-related threshold adjustments (within a DDM framework) artifactually mimicked the true effect of overall value driving choice in a stimulus-driven fashion (within an LCA framework).

We suspected that another form of model mimicry may have accounted for the other piece of evidence for motivated control in these data – the effect of value difference on decision threshold (lower thresholds for harder choices). Specifically, given that higher vs. lower value differences are associated with higher vs. lower rates of accumulation, respectively (under any of the models we've described), these decisions will reach a collapsing threshold when it is at a high vs. low point in its collapse. Thus, even though that collapsing threshold is only driven by time within a trial (i.e., it has no access to information about option value), the threshold level at which a decision is reached differs across the two cases. If a model omits this collapsing mechanism and instead assumes that thresholds are fixed within a trial – as was the case for Vassena et al.'s models [9,10] – then it may appear as though trials with lower value differences also have lower fixed thresholds. Moreover, confounding between overall value and value difference under unequal weighting (see Supplementary Figure 2) only exacerbates the potential for an artifactual influence of value difference. To test this, we again fit behavior generated by our LCA (which had a collapsing threshold) to the DDM that was consistent with the original authors' hypothesis (which had a fixed threshold). We find that this analysis mistakenly identifies a positive relationship between value difference and DDM threshold, despite there being no such connection in the generative model (Figure 3C). Thus, a model that lacks motivated control of decision threshold can not only account for behavior on this task, but can generate behavior that can be confused for motivated control.

Discussion

Decision scientists have grown increasingly interested in how we control our decision-making[1,21,22]. Building off of research into adaptive control mechanisms in other forms of cognitive tasks[23-25], studies have shown that such decision parameters can be shaped by variables such as temporal urgency[4] and choice conflict[2,11,12]. Vassena and colleagues sought to break new ground in this area by proposing a new form of value-driven adaptive control that leverages information available during a choice (option values) to guide adaptations to that choice (threshold adjustments). They provided critical model-based evidence for this account by demonstrating that difficult speeded decisions were associated with lower decision thresholds than easy ones, consistent with adaptive reductions in threshold to avoid the possibility of failing to choose either option. We reexamined their findings with a more comprehensive modeling approach, and showed that this apparent value difference-related threshold decrease arose artifactually from misspecifications in their choice model. When all features of these speeded decisions were properly accounted for, hard decisions were if anything associated with higher rather than lower thresholds. We further showed that this and other value-driven threshold adjustment predicted by this account (e.g., lower thresholds for higher-valued sets) could collectively be accounted for by bottom-up decision processes expressed in a variant on the

original modeling framework (LCA rather than DDM). Our findings thus provide alternative explanations for evidence that has been brought to bear on this novel account, while also highlighting important opportunities and pitfalls for future research into related mechanisms of adaptive control. We begin by addressing the latter before returning to the former.

Our work highlights potential pitfalls for research into the computational basis of value-based decision-making. We show that even for a relatively simple choice involving two pairs of learned values, there were at least three major modeling decisions that substantively impacted the conclusions that were ultimately drawn (Figure 4). First, we showed that model fits were improved by allowing the values in each option pair to have separate weights in the decision rather than averaging these uniformly (Figure 4A). These differential weights were evidenced in behavior, and are consistent with previous work suggesting differential attentional priority under constrained time and attentional resources[26]. Failing to account for these asymmetric weights created the false appearance that value difference and overall value were orthogonalized across choices (as originally intended), whereas accounting for these correlations reveals that these variables were in fact correlated. Thus, results previously attributed to value difference (e.g., its predicted influence on threshold) were at least in part contaminated by effects of overall value. Second, we showed that model fits were improved by allowing thresholds to collapse over time within a trial (agnostic to information about the options themselves; Figure 4B). This is consistent with a wide array of previous findings that implicate such dynamic time-based threshold adjustments[3-5], particularly in the presence of explicit time pressure as in the current study[27]. Failing to account for these collapsing bounds not only results in poorer model fits, it can also fundamentally alter one's conclusions about other parameter estimates. Specifically, we show that collapsing bounds can be mimicked by joint changes in drift rate and threshold. As a result, variables that like value difference in the present study induce increases in drift rate will also artifactually increase threshold. Covariance among these parameters can therefore also serve as one diagnostic for latent threshold dynamics. Third, we show that overall value-related variability in decision threshold within a DDM framework can mimic stimulus-driven evidence accumulation within other evidence accumulation frameworks (i.e., without positing additional top-down mechanism for threshold control; Figure 4C). While our simulations demonstrate this using an LCA, the form of mimicry we describe generalizes to any other generative choice model that naturally gives rise to overall value-related speeding effects, including other parallel accumulator models[15,28] and attentionally weighted DDMs[18].

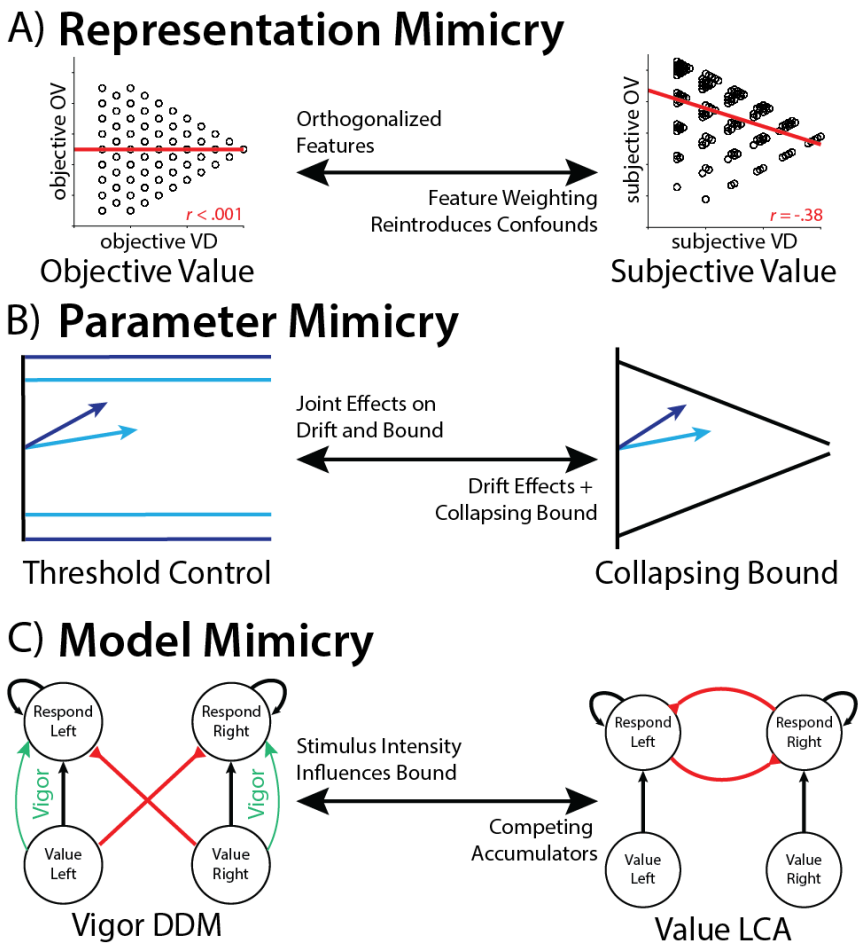


Figure 4. Multiple sources of mimicry in modeling value-based choice. A) *Representation mimicry* can occur if participants are using different information than the model assumes they use. In this experiment, participants' strategy over-weighting of high-value fractals relative to low-value fractals hid the influence of overall value on decision-making when this weighting strategy was unaccounted for. **B)** *Parameter mimicry* can occur when different parameterizations can capture similar behaviors. In this experiment, the joint effects of drift rate and threshold mimicked the influence of a collapsing bound[19]. **C)** *Model mimicry* can occur when different model classes can capture the same behavior. In this case, the influence of overall value on decision threshold could be parsimoniously accounted for by intrinsic dynamics of an independent accumulator model such as an LCA (see also [29]).

Collectively, these analyses undermine the empirical foundation for Vassena and colleagues' motivated control account. While it cannot be ruled out that future studies well lend support for this account, this empirical gap encourages further reconsideration of the *normative* foundations of this account. Most notably, the model's proposed mechanism requires that information about value (e.g., value difference) is available and used to selectively lower the decision threshold (e.g., when that difference is small). From first principles, there are problems with this requirement. First, the values of the options are precisely what the decision-maker seeks to identify as they make their decision. To the extent the decision-maker has access to these values directly, the decision process becomes obsolete. For similar reasons, existing models that account for online control adjustments do so not by relying on values per se, but instead on metacognitive representations that can be read out from the decision process (e.g., uncertainty

or conflict[2,11,12,30,31]) or from longer-run statistics of the environment (e.g., reward rate[3,32]). Second, their model predicts that, all else being equal, people will choose more quickly when choices are difficult relative to when they are easy – virtually all empirical research finds the exact opposite to be the case, both in the case of perceptual and value-based decision-making (e.g., [33-35]). Third, the model draws inspiration (and intuitive support) from applications of adaptive control to decision-making in other cognitive tasks (e.g., flanker, Stroop). However, in these tasks control can be applied to enhance the signal associated with a predetermined stimulus, location, or feature based on an explicit response rule that defines relevant and irrelevant inputs. By contrast, the response to a value-based decision – and hence which information is relevant – is inextricably tied to the information being collected online, thus making the target of control at best a moving target. The leverage that any control mechanism has for enhancing evidence or speeding up the decision process is therefore limited.

While our current findings focus on the computational mechanisms underpinning choice behavior, they nevertheless carry significant implications for understanding neural mechanisms driving such decisions. For instance, one of the aims of Vassena and colleagues' study was to test predictions of their previously proposed Predicted Response-Outcome model of dorsal anterior cingulate cortex (dACC)[36]. To provide a strong alternative to this model, they drew on a recent theory that proposes that dACC determines adjustments of cognitive control based on their overall expected value (EVC)[8,37]. The authors attributed to EVC the novel motivated control mechanism we critically re-examined here: response invigoration based on information about the values of options (Note that the original EVC model does not make this claim[19]). The authors supported the validity of this alternate model using their observation of value-based changes in response threshold, and proceeded to generate predictions about patterns of neural activity that should arise during the task. They then argued that activity in dACC during their task (greatest for the choices with the lowest value difference and the highest value difference) was inconsistent with predictions that they derived from this motivated control account, and by extension with EVC (see also [19] for additional concerns about how neural predictions were extrapolated from the EVC theory).

Our current findings raise a fundamental concern with this conclusion: their proposed motivated control account is unsupported by the findings from their experiment. Their critical finding that people selectively invigorate low value difference choices vanishes (and numerically reverses) when accounting for general choice dynamics in their task. This renders moot any neural predictions that would theoretically follow from such an adaptive control account, and may explain why Vassena and colleagues were unable to find evidence that bore out these neural predictions. It remains to be determined whether previous adaptive control accounts can quantitatively account for patterns of dACC activity observed in this study, but at a qualitative level those patterns can be accounted for by any theory that implicates the dACC in monitoring for potential conflict (low value differences) and surprise (high value differences), including EVC[1,8,19,37,38].

Substantial progress has been made toward uncovering the neural and computational mechanisms at the intersection of value-based decision-making and cognitive control. This research flourishes when bold new mechanistic accounts are proposed and tested. Recent

examples of this include work that normatively accounts for and empirically demonstrates dynamic adjustments in attention and evidence accumulation based on uncertainty and expected information gain[30,31,39], and dynamic adjustments in threshold based on conflict[2,11,12] or statistics of one's choice environment [3]. By providing a careful reexamination of both the empirical and normative basis of this recent and influential addition to these novel accounts, we hope that we have both raised the threshold for adopting such an account, while at the same time accelerating progress towards more fruitful accounts of when, whether, and how we control our decisions.

Method

Task Design. Briefly, each trial presented two fractal images on the left and right of the screen (four total). Participants had to choose between the left pair and right pair within a 750 ms deadline, or they would miss out on their choice for that trial. Each fractal was worth between €0.10 and €0.80 (in increments of €0.10), and participants received the value of one fractal from the pair they chose, chosen randomly, in addition to their participant payment. Participants performed 160 trials, which were balanced with respect to the total value and difference between the left/right options. Additional details on this task are available in the original manuscript [9].

Participants. Twenty-three participants (13 females; M_{age} = 23) were recruited, in accordance with the local research ethics committee. Twenty-two participants actually took part in the experiment.

Drift Diffusion Analysis. We fit drift diffusion models (DDMs) that both matched the authors original claims, and models that explored alternative explanations for their results. All models were fit using HDDM [13,40] , using five chains of 6000 samples (burning the first 2000 samples from each repetition). Model convergence was verified using the Gelman-Rubin R-hat statistic [41]. Apart from the Original model, models used the Likelihood approximation network (LAN) option in HDDM, providing efficient approximate likelihoods for collapsing bound DDMs. The Original model using the 'HDDM' function to predict correct responses, and the rest of the models used the 'HDDMnnRegressor' function to predict left/right choices. All models were fit with parameters at both the subject and group levels. As a requirement of the LAN fitting procedure, we used uninformed priors. Models were compared using divergence information criteria (DIC; [42]). Note that the total number of trials differs between the Original and Original* models, as the Original model removed the most difficult choices. Normalizing the deviance by the sample size ('trial-averaged likelihood') still resulted in a DIC difference of 11.1 in favor of the Original* model, with the Original* model also offering a more comprehensive account of participants' behavior. Parameter posteriors and posterior predictive checks were plotted using (Gaussian) kernel density estimation in MATLAB.

Model Name	Model Type	Regression Equation	Additional parameters
Original	ddm	v ~ [hard, easy] a ~ [hard, easy] z ~ [hard, easy]	z
Original*	ddm	v ~ -1 + signVD, a ~ 1 + absVD	z, p_outlier
VD	ddm	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD	z, p_outlier
OV	ddm	v ~ -1 + maxVD + minVD,	z, p_outlier

		a ~ 1 + maxOV + minOV	
VDOV	ddm	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD + maxOV + minOV	z, p_outlier
VD both	angle	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD, theta ~ 1 + absMaxVD + absMinVD	z, p_outlier, theta
VD init	angle	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD	z, p_outlier, theta
VD rate	angle	v ~ -1 + maxVD + minVD, theta ~ 1 + absMaxVD + absMinVD	z, p_outlier, theta
OV both	angle	v ~ -1 + maxVD + minVD, a ~ 1 + maxOV + minOV, theta ~ 1 + maxOV + minOV	z, p_outlier, theta
OV init	angle	v ~ -1 + maxVD + minVD, a ~ 1 + maxOV + minOV	z, p_outlier, theta
OV rate	angle	v ~ -1 + maxVD + minVD, theta ~ 1 + maxOV + minOV	z, p_outlier, theta
VDOV both	angle	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD + maxOV + minOV, theta ~ 1 + absMaxVD + absMinVD + maxOV + minOV	z, p_outlier, theta
VDOV init	angle	v ~ -1 + maxVD + minVD, a ~ 1 + absMaxVD + absMinVD + maxOV + minOV	z, p_outlier, theta
VDOV rate	angle	v ~ -1 + maxVD + minVD, theta ~ 1 + absMaxVD + absMinVD + maxOV + minOV	z, p_outlier, theta

Model types: ddm = static bound, angle = linear collapsing bound.
Regression Equation: v = drift rate, a = initial threshold, theta = collapse rate. signVD = signed value difference (equal weight), absVD = absolute value difference (equal weight). max/min VD = value difference for the maximum or minimum value options (e.g. max reflects the highest value left option and highest value right option). max/min OV = summed value for the max or min value options.
Additional parameters: z = starting point, p_outlier = lapse rate, theta = collapse rate.

LCA simulations. To obtain choices and response times we simulated 1000 iterations of each choice in the dataset using a custom LCA implemented in MATLAB. To account for the

overweighting of the maximum values observed in the statistical analyses (cite), we set the inputs (I) to the accumulators set to the weighted average of the two values for each option (max value weight = ⅘, min value weight = ⅕). To account for the response deadline, we implemented a linearly collapsing bound (zb) that reaches zero at the deadline. We included a standard non-decision time, t_0, which was set to 250ms. The initial bound height (z) was set to 10, the noise parameter (s) to 4.5, decay (k) to 0.153 and mutual inhibition (w) to 0.5. Accumulation was not bounded at zero, and activation could take negative values. Noise was normally distributed. Traces were simulated with a step size of 0.005. These parameters were hand-tuned to roughly approximate participants' average accuracy and reaction time to generate realistic behavioral predictions.

The accumulator activation (AA) was updates at each step as:

$$y1_{t+1} = y1_t + step*(I - k*y1_t - w*y2_t) + sqrt(step)*s(N)$$
$$y2_{t+1} = y2_t + step*(I - k*y1_t - w*y1_t) + sqrt(step)*s(N)$$

Response times were registered as the timestep when one of the accumulators met or exceeded the current bound height, and the relevant accumulator was registered as the current choice. We used a random subset of these LCA simulations to fit the 'Original*' and 'VDOV both' HDDM models described above.

Code Availability. Data is available from the original authors, all code is available at xxxx.

Acknowledgements

This research was supported by NIH grant 1R01MH124849-01, NIH grant 2P50MH106435, and NIH Shared Instrument Grant S10 OD025181, and the CV Starr Fellowship for HR. The authors are grateful to Sebastian Musslick, Jon Cohen, Matt Botvinick, and Michael Frank for valuable discussions, and to Alexander Fengler for essential advice on using the HDDM toolbox for fitting collapsing bound models.

References

1 Frömer, R. & Shenhav, A. Filling the gaps: Cognitive control as a critical lens for understanding mechanisms of value-based decision-making. *Neuroscience & Biobehavioral Reviews* **134**, 104483, doi:https://doi.org/10.1016/j.neubiorev.2021.12.006 (2022).

2 Cavanagh, J. F. & Frank, M. J. Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences* **18**, 414-421, doi:https://doi.org/10.1016/j.tics.2014.04.012 (2014).

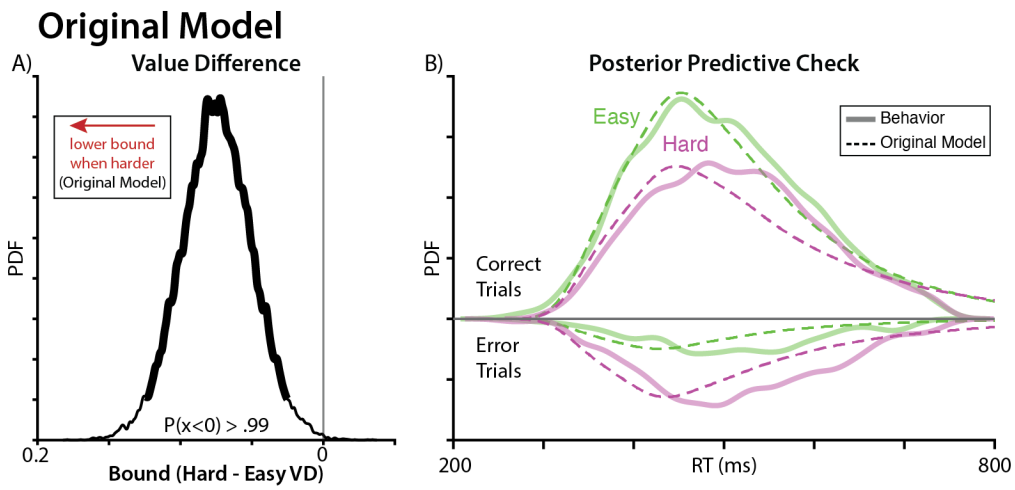
3 Tajima, S., Drugowitsch, J. & Pouget, A. Optimal policy for value-based decision-making. *Nat Commun* **7**, 12400, doi:10.1038/ncomms12400 (2016).

4 Forstmann, B. U., Ratcliff, R. & Wagenmakers, E. J. Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annual review of psychology* **67**, 641-666, doi:10.1146/annurev-psych-122414-033645 (2016).
5 Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N. & Pouget, A. The cost of accumulating evidence in perceptual decision making. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**, 3612-3628, doi:10.1523/JNEUROSCI.4010-11.2012 (2012).
6 Gluth, S., Rieskamp, J. & Buchel, C. Deciding when to decide: time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**, 10686-10698, doi:10.1523/JNEUROSCI.0727-12.2012 (2012).
7 Botvinick, M. M. Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience* **7**, 356-366, doi:10.3758/cabn.7.4.356 (2007).
8 Shenhav, A., Botvinick, M. M. & Cohen, J. D. The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron* **79**, 217-240, doi:http://dx.doi.org/10.1016/j.neuron.2013.07.007 (2013).
9 Vassena, E., Deraeve, J. & Alexander, W. H. Surprise, value and control in anterior cingulate cortex during speeded decision-making. *Nature Human Behaviour*, doi:10.1038/s41562-019-0801-5 (2020).
10 Alexander, W. & Vassena, E. Selective application of theory does not replace formal model tests: a reply to Shenhav et al., doi:https://doi.org/10.31234/osf.io/ugjrf (2020).
11 Frank, M. J. *et al.* fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **35**, 485-494, doi:10.1523/JNEUROSCI.2036-14.2015 (2015).
12 Cavanagh, J. F., Wiecki, T. V., Kochar, A. & Frank, M. J. Eye tracking and pupillometry are indicators of dissociable latent decision processes. *J Exp Psychol Gen* **143**, 1476-1488, doi:10.1037/a0035813 (2014).
13 Wiecki, T., Sofer, I. & Frank, M. HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics* **7**, doi:10.3389/fninf.2013.00014 (2013).
14 Ratcliff, R. & Starns, J. J. Modeling Confidence and Response Time in Recognition Memory. *Psychological review* **116**, 59-83, doi:10.1037/a0014086 (2009).
15 Hunt, L. T. *et al.* Mechanisms underlying cortical activity during value-guided choice. *Nature neuroscience* **15**, 470-476, S471-473, doi:10.1038/nn.3017 (2012).
16 Frömer, R., Dean Wolf, C. K. & Shenhav, A. Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications* **10**, 4926, doi:10.1038/s41467-019-12931-x (2019).
17 Sepulveda, P. *et al.* Visual attention modulates the integration of goal-relevant evidence and not value. *Elife* **9**, e60705, doi:10.7554/eLife.60705 (2020).
18 Smith, S. M. & Krajbich, I. Gaze Amplifies Value in Decision Making. *Psychological science* **30**, 116-128, doi:10.1177/0956797618810521 (2019).
19 Shenhav, A., Musslick, S., Botvinick, M. M. & Cohen, J. D. Misdirected vigor: Differentiating the control of value from the value of control. *PsyArXiv*, doi:10.31234/osf.io/5bhwe (2020).
20 Usher, M. & McClelland, J. L. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* **108**, 550-592, doi:10.1037/0033-295X.108.3.550 (2001).
21 Dayan, P. How to set the switches on this thing. *Curr Opin Neurobiol* **22**, 1068-1074, doi:10.1016/j.conb.2012.05.011 (2012).

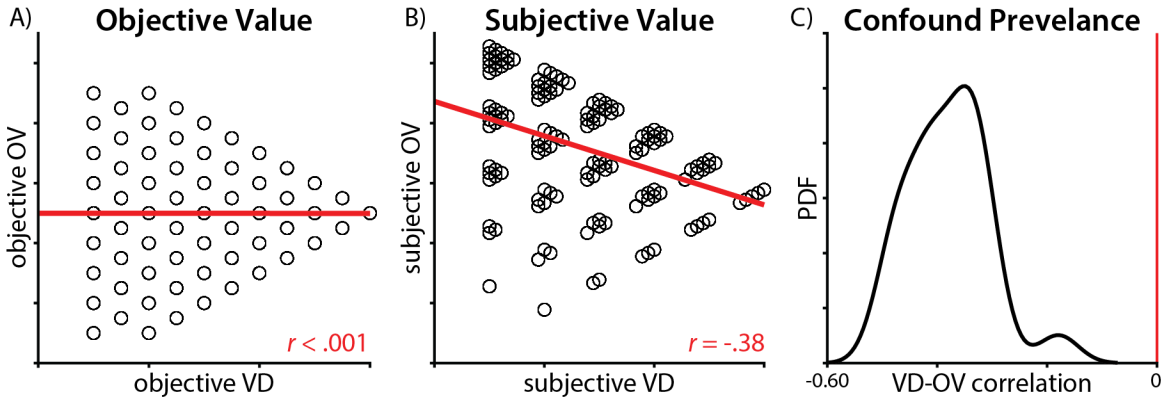
22 Simen, P. *et al.* Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *Journal of experimental psychology. Human perception and performance* **35**, 1865-1897, doi:10.1037/a0016926 (2009).
23 Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychological Review* **108**, 624-652, doi:10.1037/0033-295X.108.3.624 (2001).
24 Botvinick, M., Cohen, J. & Carter, C. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences* **8**, 539-546 (2004).
25 Yeung, N., Botvinick, M. M. & Cohen, J. D. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev* **111**, 931-959, doi:10.1037/0033-295X.111.4.939 (2004).
26 Teoh, Y. Y., Yao, Z., Cunningham, W. A. & Hutcherson, C. A. Attentional priorities drive effects of time pressure on altruistic choice. *Nat Commun* **11**, 3534, doi:10.1038/s41467-020-17326-x (2020).
27 Kelly, S. P., Corbett, E. A. & O'Connell, R. G. Neurocomputational mechanisms of prior-informed perceptual decision-making in humans. *Nat Hum Behav* **5**, 467-481, doi:10.1038/s41562-020-00967-9 (2021).
28 Wang, X. J. Neural dynamics and circuit mechanisms of decision-making. *Curr Opin Neurobiol* **22**, 1039-1046, doi:10.1016/j.conb.2012.08.006 (2012).
29 Shevlin, B. R. K., Smith, S. M., Hausfeld, J. & Krajbich, I. High-value decisions are fast and accurate, inconsistent with diminishing value sensitivity. *Proceedings of the National Academy of Sciences* **119**, e2101508119, doi:doi:10.1073/pnas.2101508119 (2022).
30 Callaway, F., Rangel, A. & Griffiths, T. L. Fixation patterns in simple choice reflect optimal information sampling. *PLoS Comput Biol* **17**, e1008863, doi:10.1371/journal.pcbi.1008863 (2021).
31 Jang, A. I., Sharma, R. & Drugowitsch, J. Optimal policy for attention-modulated decisions explains human fixation behavior. *Elife* **10**, e63436, doi:10.7554/eLife.63436 (2021).
32 Otto, A. R. & Daw, N. D. The opportunity cost of time modulates cognitive effort. *Neuropsychologia* **123**, 92-105, doi:10.1016/j.neuropsychologia.2018.05.006 (2019).
33 Oud, B. *et al.* Irrational time allocation in decision-making. *Proc Biol Sci* **283**, 20151439, doi:10.1098/rspb.2015.1439 (2016).
34 Starns, J. J. & Ratcliff, R. Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic bulletin & review* **19**, 139-145, doi:10.3758/s13423-011-0189-3 (2012).
35 Holmes, P. & Cohen, J. D. Optimality and some of its discontents: successes and shortcomings of existing models for binary decisions. *Top Cogn Sci* **6**, 258-278, doi:10.1111/tops.12084 (2014).
36 Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nature neuroscience* **14**, 1338-U1163, doi:10.1038/nn.2921 (2011).
37 Shenhav, A., Cohen, J. D. & Botvinick, M. M. Dorsal anterior cingulate cortex and the value of control. *Nature neuroscience* **19**, 1286-1291, doi:10.1038/nn.4384 (2016).
38 Shenhav, A., Straccia, M. A., Cohen, J. D. & Botvinick, M. M. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature neuroscience* **17**, 1249-1254, doi:10.1038/nn.3771 http://www.nature.com/neuro/journal/v17/n9/abs/nn.3771.html#supplementary-information (2014).
39 Frömer, R., Callaway, F., Griffiths, T. & Shenhav, A. Considering what we know and what we don't know: Expectations and Confidence guide value integration in value-based decision-making. *PsyArXiv*, doi:10.31234/osf.io/2sqyt (2022).

40 Fengler, A., Bera, K., Pedersen, M. L. & Frank, M. J. Beyond Drift Diffusion Models: Fitting a Broad Class of Decision and Reinforcement Learning Models with HDDM. *Journal of cognitive neuroscience* **34**, 1780-1805, doi:10.1162/jocn_a_01902 (2022).
41 Gelman, A. & Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* **7**, 457-472 (1992).
42 Gelman, A. *et al. Bayesian data analysis*. (Chapman and Hall/CRC, 2013).

Supplementary Figures



Supplementary Figure 1. Direct replication of analyses and findings in original paper. A) When fitting a model with the exact specifications used by Vassena et al., we replicate their finding of harder choices being associated with lower decision thresholds. **B)** Predicted choice behavior from this original model provides a poor qualitative fit to observed accuracy and reaction time distributions. See Figure 1C-D for a version of this model aimed at improving fit without altering the model's core assumptions.



Supplementary Figure 2. Unequal weighting introduces confounding between overall value and value difference. A) If all outcomes were weighted equally (as in the Original and Original* models), value difference and overall value were uncorrelated by design. **B)** If participants have unequal weighting on the best and worst outcomes in each option (which improves model fit), value difference and overall value can become correlated. Here, we used the group-level drift rate weights estimated in the best-fitting model. **C)** Using the participant-specific drift rate weights, we found that all participants show some degree of correlation between overall value and value difference, introducing the potential for confounding.