

Essential Revisions (for the authors):

Based on the combined reviews and discussions among the reviewers and review editor, the following issues should be addressed as essential revisions for moving forward.

This response letter accompanies our revised manuscript “*Flexible control of representational dynamics in a disinhibition-based model of decision making*”. In this paper, we present a new disinhibition-based circuit model of decision-making (local disinhibition decision model: LDDM), and show – via both simulations and fitting to existing empirical data - that it captures known features of the neurobiological decision process including value normalization, persistent activity, and winner-take-all choice selection. Furthermore, we show that disinhibition in this model provides a mechanism for external control of circuit state, for example switching from value coding to selection at arbitrary, task-defined times. From the original submission, the three essential revisions identified by the Reviewers and Editor were:

- The inclusion of a competing model or models (beyond models that are simply a component model of LDDM), and a conceptual discussion of what LDDM adds over existing models
- A clarification of the necessity, specificity, and interpretation of some of the model parameters in the LDDM, including an examination of optimization surfaces of the fits and parameter recovery analyses
- A more consistent conceptual framing of the goals of the paper, including a clarification of key aspects of the model that explain its behavior and a better linking of the model to previous work

We thank the Reviewers for their helpful comments, and we have revised the manuscript accordingly making all of the essential revisions specified in the reviews - we believe these changes have greatly improved the manuscript. Below we explain how we make major conceptual responses to the essential revisions; more detailed responses and specific revisions are provided in the response to specific Reviewer comments further down in this response.

Major changes in the revised manuscript are:

- The addition of an alternative competing decision model (LCA: leaky competing accumulator) for quantitative model comparison in fitting perceptual choice data and for comparative examination of best-fit model neural activity as specifically requested by the reviewers.
- New figure supplements detailing optimization surfaces for model fits as requested.
- New figure supplements showing the reviewer-suggested parameter recovery analyses.
- Significant revisions to the Introduction and Discussion to clarify the conceptual framing of the paper in the manner suggested by the reviewers; the new framing emphasizes the theoretical and empirical importance of incorporating disinhibition into circuit models of decision-making, and utilizes the integration of value normalization and WTA activity as a conceptual testing of the LDDM rather than an end goal in and of itself
- A new discussion of existing disinhibition-based decision models in the cortical basal ganglia (CBG) system, acknowledging this related prior work as requested, this includes an overview of similarities and differences between these existing models and the LDDM
- A new discussion of the relationship between the LDDM and previous computational models that have employed disinhibition
- Revision of model figures (**Figs. 2, 2-figure supplement1A**) to better explain the unique features we refer to as disinhibitory control (as a parameter in the disinhibitory *D* neuron that controls responsiveness to excitatory *R* neuron input)

1) Competing models: While all three reviewers agree that a network that combines both value-based and WTA dynamics is interesting and useful, there is consensus that the lack of competing or contrastive models (beyond the component models that are combined to make the LDDM) tempers the conclusions

that can be taken away from the work. The LDDM should be compared with reasonable competing models and, conceptually, the authors should highlight what the LDDM adds over existing models).

In the revision, we now also compare LDDM performance with an additional standard circuit model of decision-making - the leaky competing accumulator model (LCA). Model comparison shows that the LDDM outperformed both the RNM and the LCA in fitting a standard decision-making dataset. We would like to point out that the RNM has been prevalent for two decades, and we believe it is appropriate to be considered as one of the standard decision-making circuit models. However, the LCA is a widely known alternative dynamical model of decision-making suitable for behavioral comparison. Thus, we have now added the LCA as model fits to monkey WTA behavior, allowing a comparison across the three models (LDDM, RNM, LCA); in addition, we also quantify and compare how the option-coding unit activity in each best-fit model matches neurophysiological data.

Based on the Reviewers' suggestions, we also now highlight two fundamental differences between the LDDM and other existing circuit models. First, we now highlight that the common inhibitory motif in existing standard decision-making models (e.g., RNM) is non-selective inhibition. We clarify here in the revision that the LDDM has a fundamental difference from RNM in predicting selective (or structured) inhibition, which has been identified in recent empirical studies using advanced neural imaging techniques. Second, we now explain that disinhibition in the LDDM provides a circuit mechanism for a switch between valuation and WTA dynamics; such a mechanism does not exist in simpler models like the RNM. Thus, while we show that the LDDM can outperform alternative models quantitatively, we now highlight that the LDDM is a novel framework that is qualitatively different from previous models, can accommodate new empirical findings (e.g., selective inhibition), and predicts specific hypotheses (e.g., differential activity of different inhibitory interneuron subtypes) for testing in future studies.

2) Parameter specificity: There is a consensus in the reviews regarding questions about the necessity, specificity, and interpretation of some of the model parameters. Reviewer 3, in particular, points out potential inconsistencies in the way the disinhibition Beta parameter is interpreted. Reviewer 2 highlights confusion between the necessary and specific role of Alpha compared to Beta in some of their simulations. As both reviewers point out, an examination of the optimization surface of the fits (Reviewer 3) and/or a parameter recovery analysis (Reviewer 2) could help demonstrate the robustness and identifiability of the LDDM model parameters. This would also address Reviewer 1's concern about model complexity and overfitting.

We thank the Reviewers for pointing out inconsistencies in notation regarding the β parameter. In the revision, we have fixed the notation problem on the β parameter (sorry about that) such that it is consistent throughout the manuscript. The manuscript now clarifies that in the LDDM the β parameter controls D unit response to R unit activity, and we now highlight that we conceptualize it as a measure of the functional connectivity between R and D neurons.

Regarding the role of α and β , we clarify in the response below and in the revision the uniqueness of α and β contributions to the dynamics of the system. The main takeaway, now highlighted in the text, is that for WTA selection dynamics in the LDDM, β is required while α contributes but is not necessary. In the revision, we now further examine and visualize the optimization surface of the fits for all parameters (as requested). Most of the parameters show smooth likelihood spaces, a narrow optimization range, and small collinearities, which indicate good identification of the model parameters and parsimonious model complexity. The α parameter shows less precise identification and shares some collinearity with β . We examined this problem carefully and realized that the α and β parameters make differential contributions to the shape of reaction time distribution in perceptual choice; these relatively small but potentially important differences in the likelihood space is an interesting phenomenon that we plan on targeting in future studies. Overall, after careful examination on the model fitting including parameter recovery analyses, we find that the model fitting is reliable and the parameters settings are parsimonious. These new analyses are provided in a series of additional figures and additional text (see below for specifics).

3) Conceptual framing: The reviews point out that the conceptual framing of the goals shifts across the different sections of the manuscript. The authors should be clear about the high-level framing (Reviewer 1), whether it's about the integration of two decision frameworks or the role of disinhibition). The author should also clarify the precise interpretation of the key aspects of the model that explain its behavior (Reviewers 1 & 3). Finally, the authors should more extensively link their model in the context of prior work (Reviewers 1 & 3)

The issue of conceptual framing is an important one, and we thank the Reviewers for pointing out the changing framing in the original paper. We have now edited the paper to highlight the importance of inhibition in computational models of decision-making, focusing throughout on two issues in the existing literature: (1) the assumption of non-selective, pooled inhibition in decision models (which is not supported by recent empirical evidence), and (2) widespread empirical evidence for interneuron diversity and complexity in inhibition circuits. Given this framing, the paper is now framed by introducing disinhibition into a circuit model of decision-making. In this framing, the integration of normalized value coding and WTA activity remains essential as a test of the validity of the LDDM.

Regarding the interpretation of model features, we have revised the paper to emphasize the relative importance of the β and α parameters to LDDM behavior, the conceptual role of β as a functional coupling between excitatory and disinhibitory units, and the interpretability of fit α values in terms of anatomical recurrence. Finally, we have now also added significant new text to the Discussion to place the LDDM in the proper context of previous computational work. As suggested by Reviewer #1, this includes a discussion of models of action selection in cortical-basal ganglia loops, which also utilized disinhibition (though with some differences that the LDDM, this is now discussed explicitly). As suggested by Reviewer #3, we also discuss how the LDDM is related to other computational models that use disinhibitory circuit motifs and address the functional processes relevant to the LDDM (divisive normalization, working memory, and decision-making).

eLife assessment

This work provides a promising first pass at providing an integrative model for how decisions arise from neural circuits. The approach is novel but lacks a more rigorous vetting against alternative model formulations to be able to determine its true significance. More stringent evaluations of the model in the context of existing work, as well as a clearer description of the goals and implementation of the approach, would help to address these concerns.

Reviewer #1 (Public Review):

This work presents a unification model (of sorts) for explaining how the flow of evidence through networks can be controlled during decision-making. The authors combine two general frameworks previously used as neural models of cortical decision-making, dynamic normalization (that implement value encoding via firing activity) and recurrent network models (which capture winner-take-all selection processes) into a unified model called the local disinhibition-based decision model (LDDM). The simple motif of the LDDM allows for the disinhibition of excitatory cells that represent the engagement of individual actions that happens through a recurrent inhibitory loop (i.e., a leaky competing accumulator). The authors show how the LDDM works effectively well at explaining both decision dynamics and the properties of cortical cells during perceptual decision-making tasks.

All in all, I thought this was an interesting study with an ambitious goal. But like any good study, there are some open issues worth noting and correcting.

MAJOR CONCERNS

1. Big picture

This was a comprehensive and extremely well-vetted set of theoretical experiments. However, the scope and complexity also made the take-home message hard to discern. The abstract and most of the introduction focus on the framing of LDDM as a hybrid of dynamic normalization models (DNM) and recurrent network models (RNMs). This is sold as a unification of value normalization and selection into a novel unified framework. Then the focus shifts to the role of disinhibition in decision-making. Then in the Discussion, the goal is stated as to determine whether the LDDM generates persistent activity and does this activity differ from RNMs. As a reader, it seems like the paper jumps between two high-level goals: 1) the unification of DNM and RNM architectures, and 2) the role of disinhibition. This constant changing makes it hard to focus as the reader goes on. So what is the big picture goal specifically?

Also, the framing of value normalization and WTA as a novel computational goal is a bit odd as this is a major focus of the field of reinforcement learning (both abstractly at the computational level and more concretely in models of the circuits that regulate it). I know that the authors do not think they are the first to unify value judgements with selection criteria. The writing just comes across that way and should be clarified.

We thank the Reviewer for their thoughtful consideration of the overall framing of the big picture goals of the paper. Upon reflection, we agree that the paper really centers on the importance of incorporating disinhibition into computational circuit-based models of decision-making. Thus, we have significantly revised the Introduction and Discussion to focus on the theoretical and empirical importance of incorporating disinhibition into computational models of decision-making, and use the integration of value normalization and WTA selection as an example of how disinhibition increases the richness of circuit decision models. Please see the response to recommendations below for more detail on the changes.

2. Link to other models

The LDDM is described as a novel unification of value normalization and winner-take-all (WTA) selection, combining value processing and selection. While the authors do an excellent job of referencing a significant chunk of the decision neuroscience literature (160 references!) the motif they end up designing has a highly similar structure to a well-known neural circuit linked to decision-making: the cortico-basal ganglia pathways. Extensive work over the past 20+ years has highlighted how cortical-basal ganglia loops work via disinhibition of cortical decision units in a similar way as the LDDM (see the work by Michael Frank, Wei Wei, Jonathan Rubin, Fred Hamker, Rafal Bogacz, and many others). It was surprising to not see this link brought up in the paper as most of the framing was on the possibility of the LDDM representing cortical motifs, yet as far as I know, there does not exist evidence for such architectures in the cortex, but there is in these cortical-basal ganglia systems.

We thank the Reviewer for the suggestion to link the LDDM to disinhibition in CBG models; this is indeed an important body of empirical and computational work that we overlooked in the original manuscript. We have now added text to the Discussion to highlight the link between LDDM and these CBL disinhibition models, focusing on how they are conceptually similar and how they differ. Please see our response to recommendations below for a more detailed discussion of the revisions.

3. Model evaluations

The authors do a great job of extensively probing the LDDM under different conditions and against some empirical data. However, most of the time there is no "control" model or current state-of-the-art model that the LDDM is being compared against. In a few of the simulation experiments, the LDDM is compared against the DNM and RNM alone, so as to show how the two components of the LDDM motif compare against the holistic model itself. But this component model comparison is inconsistently used across simulation experiments.

Also, it is worth asking whether the DNM and RNM are appropriate comparison models to vet the LDDM against for two reasons. First, these are the components of the full LDDM. So these tests show us how the two underlying architectural systems that go into LDDM perform independently, but not necessarily how the LDDM compares against other architectures without these features. Second, as pointed out in my previous comment, the LDDM is a more complex model, with more parameters, than either the DNM or RNM. The field of decision neuroscience is awash in competing decision models (including probabilistic attractor models, non-recurrent integrators, etc.). If we really want to understand the utility of the LDDM, it would be good to know how it performs against similarly complex models, as opposed to its two underlying component models.

We greatly appreciate the Reviewer's comments on the point of model comparison, which points out that our original manuscript failed to clearly convey a very important difference between the LDDM and the existing RNM(s). In the revision, we now make it clearer that the fundamental difference between the LDDM and the RNMs is the architecture of disinhibition (see the revised Introduction, especially p. 8 lines 164-168). The LDDM is not simply a combination of the DNM model with RNM architecture (a point we may have mistakenly conveyed in the original manuscript): the introduction of disinhibition separates LDDM inhibition into option-selective subpopulations, as opposed to the single pooled inhibition of RNM models. Given this fact, the LDDM predicts unique selective-inhibition dynamics shown in recent optogenetic and calcium imaging results, a finding inconsistent with the common-pooled and non-selective inhibition assumed in the existing RNMs and many of its variants. Thus, we believe that a comparison between the LDDM and the RNM, which share similar level of complexity and numbers of parameters, is important.

We also appreciated the Reviewer's concern about testing the LDDM against alternative models. In order to better connect to the existing literature, we now compare the LDDM to another standard circuit model of decision-making - the leaky competing accumulator (LCA) model. The LCA is a circuit model that captures many of the aspects of perceptual decision-making seen in the mathematical drift diffusion model (DDM), but with a construction that allows for fitting to behavioral data and comparison of underlying unit activities. Please see our response to recommendations below for further detail.

4. Comparison to physiological data

I quite enjoyed the comparisons of the excitatory cell activity to empirical data from the Shadlen lab experiments. However, these were largely qualitative in nature. In conjunction with my prior point on the models that the LDDM is being compared against, it would be ideal to have a direct measure of model fits that can be used to compare the performance of different competing "control" models. These measures would have to account for differences in model complexity (e.g., AIC or BIC), but such an analysis would help the reader understand the utility of the LDDM in connecting with empirical data much better.

We agree with the Reviewer that a quantitative comparison of the match between model neural predictions and empirical neurophysiological data is important. First, we wish to clarify that the model neural predictions are simulated from models fit to the behavioral (choice and RT data), not from fits to the neural activity traces – a point we now clarify in the text. While directly fitting dynamic models (LDDM, RNM, or LCA) to the neurophysiological data is appealing, there are currently several obstacles to this approach. The first problem is the complexity of the dynamic neural traces. Despite the long history of the random-dot motion paradigm, detailed features of the

dynamics are still not understood. For example, the stereotyped initial dip after stimulus onset may reflect a reset of the network state to improve signal to noise ratio (Conen and Padoa-Schioppa, 2015) or simply reflect a surround suppression-like lateral inhibition in visual processing. A second problem is that the primary difference between the models is the activity of inhibitory (and disinhibitory) neurons, which are typically not recorded in neurophysiological experiments; thus, there is a lack of empirical data to which to fit the models. In the revision, we clarified that the model fitting to the Roitman & Shadlen data is for behavioral data only, and model unit activity traces are derived from models fit to behavioral data.

That being said, we agree that a quantitative comparison of model activity predictions is helpful. Because the models are fit not to the neural data but to the behavioral data, rather than using likelihood-based measures like AIC and BIC we used a simple RMSE measure to compare the match between predicted and neural activity patterns (revised Fig. 6E, Fig 6-S4E, Fig 6-S5E). Please see response to recommendations below for details.

Reviewer #1 (Recommendations for the authors): Numbers refer to each point made in the Public Review.

1. I recommend the authors take time to more explicitly clarify the goal of the study. What is the singular take-home message that the reader should take away from this? This singular message should be tempered enough so as not to overstate this as the first unification of value normalization and response selection, but more specific to what is being tested.

Upon reflection, we agree with Reviewer #1 that the original manuscript was not sufficiently clear in its emphasis, touching on a number of features of the model and findings. Ultimately, we believe that the most important aspect of the LDDM is disinhibition, and the most salient contribution of this paper is showing the advantages of a decision-making circuit that incorporates disinhibition (along with known features such as recurrent excitation and lateral and/or divisive inhibition). Driven by advances in genetic and imaging technologies, recent experiments have examined the diversity of inhibitory neuron subtypes and interneuron-interneuron connectivity, revealing evidence for widespread circuits implementing network disinhibition and driving behavioral and cognitive functions. However, disinhibition is not a part of standard circuit models of decision-making, and the potential contributions of disinhibition to neurophysiological and behavioral aspects of decision-making are unknown.

Thus, the overarching goal of the paper is to design, implement, and characterize the behavior of a decision-making circuit with disinhibitory motifs. The unification of value normalization and WTA activity is important, but primarily as a demonstration of the validity of the disinhibition-based LDDM model. As demonstrated in the paper, there are additional important capabilities of the LDDM, including the generation of structured inhibition and the provision of a mechanism for top-down control of circuit dynamics and function.

To make the take-home message clearer, we have now substantially revised the Abstract to emphasize the goal of incorporating disinhibition into a dynamical circuit model of decision-making:

“Inhibition is crucial for brain function, regulating network activity by balancing excitation and implementing gain control. Recent evidence suggests that beyond simply inhibiting excitatory activity, inhibitory neurons can also shape circuit function through disinhibition. While disinhibitory circuit motifs have been implicated in cognitive processes including learning, attentional selection, and input gating, the role of disinhibition is largely unexplored in the study of decision-making. Here, we show that disinhibition provides a simple circuit motif for fast, dynamic control of network state and function. This dynamic control allows a novel disinhibition-based decision model to reproduce both value normalization and winner-take-all dynamics, the two central features of neurobiological decision-making captured in separate existing models with distinct circuit motifs. In addition, the disinhibition model exhibits flexible attractor dynamics consistent with different forms of persistent activity seen in working memory. Fitting the model to empirical data shows it captures well both the neurophysiological dynamics of value coding and psychometric choice behavior. Furthermore, the

biological basis of disinhibition provides a simple mechanism for flexible top-down control of network states, enabling the circuit to capture diverse task-dependent neural dynamics. These results suggest a new biologically plausible mechanism for decision-making and emphasize the importance of local disinhibition in neural processing.”

In addition, we have restructured the Introduction text to frame the importance of incorporating disinhibition into computational decision models. The primary change is to the initial paragraphs, though we have streamlined the Introduction as well:

“Inhibition is an essential component in neural network models of decision-making. In standard decision models, pools of option-selective excitatory neurons compete in a winner-take-all selection process via feedback inhibition (Roach et al., 2023; X.-J. Wang, 2002; Wong & Wang, 2006). Generally, such inhibition is thought to be homogenous and non-selective, with a single pool of inhibitory neurons receiving broad excitation, and in turn inhibiting excitatory neurons. However, more recent empirical findings suggest that inhibitory neurons interact with the decision circuit in a more structured manner. Inhibitory neurons active in decision-making exhibit choice-selective activity on par with excitatory neurons in the frontal cortex (Allen et al., 2017), parietal cortex (Allen et al., 2017; Najafi et al., 2020), and striatum (Gage et al., 2010), in contrast to the non-selective or broadly tuned inhibition seen in visual cortex during stimulus representation (Bock et al., 2011; Chen et al., 2013; Hofer et al., 2011; Kerlin et al., 2010; Liu et al., 2009; Niell & Stryker, 2008; Sohya et al., 2007). At an anatomic level, inhibitory interneurons also exhibit a remarkable diversity in morphology, connectivity, and physiological functions (Kepecs & Fishell, 2014; Markram et al., 2004; Tremblay et al., 2016). A prominent circuit motif observed in these anatomical studies is local disinhibition, in which vasoactive intestinal peptide (VIP)-expressing interneurons inhibit the neighboring interneurons expressing somatostatin (SST) or parvalbumin (PV) that inhibit dendritic or perisomatic areas in pyramidal neurons, so that locally disinhibit the activities of the pyramidal neurons in the neighboring area (Chiu et al., 2013; Fino & Yuste, 2011; Fu et al., 2014; Karnani et al., 2014, 2016; S. Lee et al., 2013; Letzkus et al., 2011; Pfeffer et al., 2013; Pi et al., 2013; Urban-Ciecko & Barth, 2016). Here we explore the computational implications of that motif in decision-making.

While disinhibitory circuit motifs have been implicated in cognitive processes including learning, attentional selection, and input gating (Fu et al., 2014; Letzkus et al., 2011; X.-J. Wang & Yang, 2018), how disinhibition functions in decision-making circuits is unknown. Local circuit inputs to the VIP neurons suggest that disinhibition may be a key mechanism for generating the mutual competition necessary for option selection in decision-making. In addition, given the existence of long-range inputs (Kepecs & Fishell, 2014; S. Lee et al., 2013; Pfeffer et al., 2013; Pi et al., 2013; Schuman et al., 2021) and neuromodulatory inputs (Alitto & Dan, 2013; Fu et al., 2014; Pfeffer et al., 2013; Prönneke et al., 2020; Rudy et al., 2011; Tremblay et al., 2016) to the VIP neurons, local disinhibition has been proposed to play a particular role in dynamic gating of circuit activity; such gating may be essential in decision circuits underlying flexible behavior, mediating top-down control of network function (Fu et al., 2014; Kamigaki, 2019; S. Lee et al., 2013; Letzkus et al., 2011; Pi et al., 2013; Schuman et al., 2021; S. Zhang et al., 2014). Here we hypothesize that disinhibition controls a transition between information processing states, allowing a single decision-making circuit to both represent the values of alternatives and select a single best option amongst those alternatives.”

We have also added additional framing sentences in the Discussion (p. 32, lines 684 - 689):

“The prevalence of disinhibitory circuit motifs in the brain, and recent evidence for structured decision-related inhibitory activity, argue for a more structured implementation of inhibition in computational models of decision-making than has been previously employed. Here, we show that the disinhibition-based LDDM replicates three characteristic features of observed neurobiological decision-making circuits – normalized value coding, WTA choice, and persistent activity – for the first time within a single circuit architecture.”

2. I recommend adding a discussion on known disinhibition circuits like the cortical-basal ganglia loops and showing how the LDDM links to prior models of these networks.

The reviewer rightly points out an existing and important body of work that examines disinhibition circuits functioning in motor selection (and perhaps other types of selection), specifically cortical-basal ganglia loops (CBG). Disinhibition has multiple roles in the pathways that comprise CBG circuitry, the most relevant being the direct pathway where cortical activation of striatal GABAergic medium spiny neurons inhibits tonically active GPi/SNr inhibitory neurons, thus releasing downstream neurons in the output pathway (thalamus). LDDM is related to precursor CBG models because, in both cases, activation of disinhibition is part of the selection process. This important historical link is now acknowledged in the manuscript.

A crucial difference between disinhibition in CBG and LDDM is the selectivity of the disinhibition and its role in specifying versus initiating choice. In standard models of the CBG, disinhibition is selective and favors the option to be chosen; this selective or biased disinhibition is driven by differences in input (in simple selection models) or differences in synaptic weighting in the striatum (in reinforcement learning models of the BG). In contrast, initial disinhibition in the LDDM is broad and non-selective, activated across all option subcircuits (e.g., via broadcasting projection of acetylcholine or serotonin) and serving to transform circuit function from normalized value coding to WTA selection. Subsequently, because disinhibitory neurons are driven by local excitatory (*R* neuron) input and interact in a multiplicative fashion, disinhibition becomes selective and biased, similar to CBG models. A more subtle difference between the models lies in the structure of mutual competition: like other cortical decision models (i.e. RNM), the LDDM utilizes lateral inhibition to provide competition between option-selective neurons; lacking clear evidence for such lateral connections in the BG, CBG models achieve competition in a more complicated manner that typically involves direct pathway disinhibition coupled with broad inhibition via the indirect and/or hyper direct pathways (for example, involving the subthalamic nucleus). This novel aspect of the LDDM, with regard to CBG models, is now highlighted in the manuscript.

We have now added additional text to the Discussion section (pp. 36-37) describing known disinhibition circuits in the CBG, discussing the link between the LDDM and previous computational models, and touching on the novel contribution of the LDDM:

“While largely absent in standard existing cortical decision models, disinhibition is a key element of action selection in models of the cortical-basal ganglia (CBG) system (Bogacz & Gurney, 2007; Frank, 2005; Lo & Wang, 2006; Schroll & Hamker, 2013; Wei et al., 2015). In the basal ganglia direct pathway, GABAergic neurons in the striatum inhibit neurons in the substantia nigra pars reticulata and internal globus pallidus, which in turn send inhibitory projections to the thalamus. Cortical inputs to the striatum thus produce a disinhibition of thalamic outputs to the cortex and brainstem motor areas, resulting in motor facilitation. Crucially, the activation of disinhibition in the CBG system is selective: the selection of a specific action requires a selective disinhibition driven by asymmetries in cortical inputs or striatal synaptic weights. This selective disinhibition is an essential element of computational models of the CBG system (Frank, 2005; Lo & Wang, 2006), including more complex models that incorporate global inhibition mediated by the indirect and hyper-direct pathways (Bogacz & Gurney, 2007; Schroll & Hamker, 2013; Wei et al., 2015).

While both the LDDM and standard CBG models utilize disinhibition to drive selection, they differ in two important ways. First, disinhibition in the LDDM specifically functions to implement a transition between value coding and WTA selection states. This transition is mediated by a broad/non-selective activation of disinhibition across the decision circuit. The activation of disinhibition is not biased towards specific alternatives until a period of interaction with differential value inputs to option-specific subcircuits that instantiates the WTA process. Second, disinhibition in the LDDM is tightly integrated with the lateral inhibition that mediates competition (and hence normalization) between alternatives; consistent with the microarchitecture of the cortex which it seeks to model (Fu et al., 2014; Karnani et al., 2016; Kepecs & Fishell, 2014; Pi et al., 2013; S. Zhang et al., 2014), disinhibitory, inhibitory and excitatory neurons are part of the same local circuit. In contrast, the basal ganglia are known to lack these local, lateral connections and mutual competition. As a result CBG models typically require both direct pathway disinhibition along

with diffusive suppression of competing motor plans via the indirect or hyper-direct pathways (Bogacz & Gurney, 2007; Schroll & Hamker, 2013; Wei et al., 2015) for effective operation. Thus, while conceptually similar to the CBG models, disinhibition in the LDDM is in some ways quite distinct, tightly integrated with competitive inhibition, and providing dynamic control of circuit state, both characteristics of decision-making in cortical brain areas.”

3. I would recommend finding a non-DNM and non-RNM control model to compare the LDDM against.

We appreciate the suggestion to compare the performance of the LDDM against another model, particularly in terms of WTA activity. First, we wish to clarify a point of interpretation that we did not express clearly in the original manuscript. The RNM is not strictly speaking a component of the LDDM: a key feature of RNM models is they implement a pooled inhibition, in which different option-specific excitatory neurons receive the identical inhibition signal; in contrast, the LDDM has a more complicated inhibitory structure, since the disinhibition - once activated - is driven by local excitatory neuron activity and thus functionally segregates the inhibitory pools. At a practical level, RNM models predict non-selective inhibition that does not encode any decision-relevant information; in contrast, the LDDM predicts selective inhibition where the activity of a given G neuron will reflect choice-related information relevant to its associated R neuron. Importantly, recent empirical evidence suggests that inhibition in decision circuits is selective and structured rather than non-selective. Our revision makes this issue much clearer and focused on this novel feature of the LDDM.

However, we do appreciate that including an additional model for comparison is informative and helpful and so have added a fourth model to the paper. While RNM models are by far the most prevalent model of decision-making circuits, another well-known model is the leaky competing accumulator (LCA) model (originally proposed by Usher and McClelland, 2001). The LCA is an appropriate choice for model comparison here because it is: (1) a biologically-inspired model of choice, (2) specifically designed to capture choice data in perceptual decision tasks (such as used in the Roitman and Shadlen dataset), and (3) intended to capture both psychometric (choice) and chronometric (RT) aspects of relevant decision data.

The LCA model is now described in the Methods section (pp. 51-52):

“Fitting the LCA to empirical behavioral data

Another widely acknowledged decision circuit model – the leaky competing accumulator model (LCA) (Usher & McClelland, 2001) was fit to the behavioral data (Roitman & Shadlen, 2002). The dynamics of the two nodes in the LCA can be described using the following differential equations (Eq. 22).

$$dx_i = \left(\rho_i - kx_i - \beta \sum_{j \neq i} x_j \right) \frac{dt}{\tau} + \xi_i \sqrt{\frac{dt}{\tau}} \quad (22)$$

where x_i ($i = 1$ and 2) indicates the activity of each node; ρ_i indicates the excitatory input value to each node; k indicates the net leakage on each node after the cancellation of recurrent excitation; β weighs the mutual inhibition strength from the other nodes; ξ_i is a Gaussian random noise on each node with a standard deviation of σ .

The input values ρ_i were set as $1+c'$ for Option 1 and $1-c'$ for Option 2, with c' changing over 0 to .512. We fitted the threshold as a free parameter. In that way, the time constant τ can be taken as an arbitrary value (100 ms used in our case) since it was not independent from the threshold. Other than the parameters we mentioned above, non-decision time T_0 was fixed as 120 ms, sharing the same assumption with the other two models based on the empirical observed delays after stimulus onset (90 ms) and before saccade (30 ms). That gives in a total of four free

parameters to estimate k , β , σ , and *threshold*. Since the scale of the activities is arbitrarily defined, it would need rescaling when compared to the empirical data of mean firing rates in the unit of Hz. The task setting and the optimization used were kept the same as in fitting the LDDM (see above). The time step dt was set as .001 s.”

Results of the model comparison between the LDDM, RNM, and LCA models are now described in the Results, and the fitted LCA data are shown in a new Figure 6-figure supplement 5. The following text has been added to the Results section (p. 20, lines 424-432):

“We compared the performance of the LDDM in fitting this classical dataset with the reduced form of the RNM (Wong & Wang, 2006) (**Fig. 6-figure supplement 4**), as well as another prominent computational decision model with a similar architecture of mutual inhibition – the leaky competing accumulator model (LCA; Usher & McClelland, 2001; see **Fig. 6-figure supplement 5**). The performances of the three models were close in predicting averaged RTs and choice accuracy (panel **C**). However, the LDDM captures the skewness and the shape of RT distributions better than the other two, as reflected in goodness of fit (negative log-likelihood) and AIC measures ($nLL_{LDDM} = 16546$, $nLL_{RNM} = 16573$, $nLL_{LCA} = 16948$, $AIC_{LDDM} = 33109$, $AIC_{RNM} = 33165$, $AIC_{LCA} = 33932$).”

4. I recommend using model fit metrics to evaluate how well the LDDM (and a control model) explain the neurophysiological data.

We appreciate the Reviewer’s constructive recommendation. As outlined in the response to the public review above, we now clarify that we examined model *neural* activity in three different models (LDDM, RNM, and now LCA) best-fit to behavioral data. Adequately fitting the model(s) to neurophysiological data – particularly the fully dynamics of decision-related activity – will require more work and empirical observations, and given current technical limitations we do not address fitting the models to neural data in the current manuscript.

However, we believe there are informative aspects of the approach we currently take – fitting the models to behavioral (choice and RT) data, and then examining the neural activity of the best fitting model units. We now quantify how well different model excitatory unit activity – with model parameters fit to behavior – matches empirically recorded data (from Roitman & Shadlen), focusing on how population average activity varies with stimulus level (coherence) at different timepoints in the trial. Because these models were not fit to the neural data, we use simple RMSE as a measure of explanatory power.

This information is now presented in Fig. 6E, 6-S4E, and 6-S5E, along with new Results text (pp. 20-21, lines 443-448 and lines lines 450-461):

“More quantitatively, we examined the relationship between activity and coherence at the specific time point reported in the original work (arrow points **a** and **b**, **Fig. 6E**). Model predictions align well with empirical observations: across the three alternative models, the deviation between empirical recordings and model-predicted activity is the smallest for LDDM (quantified by root-mean-square error (RMSE); $RMSE_{LDDM} = 2.74$ (**Fig. 6E**), $RMSE_{RNM} = 20.10$ (**Fig. 6-figure supplement 4E**), $RMSE_{LCA} = 3.92$ (**Fig. 6-figure supplement 5E**)).

Aligned to the onset of decision (**Fig. 6D**, right), ... Quantification shows that LDDM again best predicted empirical neural activity with data aligned to choice onset ($RMSE_{LDDM} = 6.77$ (**Fig. 6E**); $RMSE_{RNM} = 9.35$ (**Fig. 6-figure supplement 4E**); $RMSE_{LCA} = 7.51$ (**Fig. 6-figure supplement 5E**)). Thus, *R* unit activity – in a model with parameters fit only to behavior – replicates the recorded activity of parietal neurons during both initial decision processing and eventual choice selection.”

On a broader note, a significant limitation of fitting neural data and using model comparison metrics is that the novel focus of the model - the activity of inhibitory and disinhibitory neurons - has not been empirically well quantified in these types of tasks. Vary little data from these classes of neurons is currently available. In the revised manuscript,

we address this limitation by describing model predictions about the dynamics of different neuronal types and the perturbation outcomes to guide future data collection. We hope that such model predictions, along with future work recording from identified neural subpopulations, will help deepen the understanding of circuit mechanisms underlying the decision-making process.

To highlight this issue we now visualize the model predictions regarding the dynamics of the inhibitory units (*G*) and disinhibitory units (*D*) for future empirical testing (please see revised Figure 6F-I). We have also included a description of the informative pattern in the predicted dynamics of *G* and *D* in Results (pp. 21-22).

“Unlike the RNM and LCA models, the LDDM predicts different dynamics in different subtypes of interneurons (**Fig. 6F-I**). The inhibitory (*G*) units selectively code input values and choice but exhibit complex dynamics due to the interplay of feedforward excitation, lateral inputs, and disinhibition. Early on (dynamics sorted to the left in **Fig. 6F** and upper panel in **Fig. 6G**), the *G* activities initially increase due to excitatory drive from *R* units. Later on, when the inhibition from *D* units increases (**Fig. 6H**), the *G* activities start to decrease. Near the time of choice (dynamics sorted to the right in **Fig. 6F** and the lower panel in **Fig. 6G**), the chosen *G* units show lower activities than the unchosen side because of stronger inhibition from *D* as an outcome of WTA competition. The dynamics of *D* units rapidly increase in the early stage, driven by excitatory *R* unit activity (dynamics sorted to the left in **Fig. 6H**). Dynamics in the late stage (dynamics sorted to the right in **Fig. 6H**) show higher activity on the chosen side than the unchosen side as an outcome of WTA competition. Both types of interneurons show different time-dependent patterns of coherence-dependence that likely reflect the complex dynamics of the system and RT-based data aggregation methods (**Fig. 6G, H**). While the activity of different interneuron subtypes have not been widely recorded in decision tasks, these new LDDM predictions provide a testbed for future empirical and theoretical investigations.”

MINOR CONCERNS

A. How were the weights determined? When comparing the LDDM to the RNM/DNM, were the same weights used, or was the network retuned?

In the model fittings, the connection weights of different models were independently fit to achieve the best performance of every model in capturing the data. In fitting to the neurophysiological data of value representation (Louie et al., 2011), two, three, and four free parameters were assumed for DNM, LDDM, and RNM, respectively. In fitting to the winner-take-all binary choice data (Roitman & Shadlen, 2006), 7 and 8 parameters were assumed for LDDM and RNM, respectively. Different numbers of parameters used in fitting different datasets is due the limitation of each dataset in testing different aspects of functions. For example, the dataset of Louie et al. (2011) is restricted to test the function of value representation but not choice. Parameters related to choice are not able to be specified in this dataset.

B. The attractor analysis was incredibly interesting and powerful. However, the description of it in the Results is light, especially if the authors are hoping to appeal to a broad audience. In particular, the description in lines 195-204 needs to be significantly expanded in order to walk the reader through the phase plots in Fig. 3D.

In the revision, we have elaborated the description of the phase plane analysis related to Fig. 3D to better walk through our approach to a broad audience. We agree that a better conceptual understanding of the attractor implications is essential for a reader.

In Results (pp. 12, lines 240 -256), we have now revised the text describing the phase plane analysis to read:

“Taking advantage of its simplified mathematical form, we analytically evaluated the LDDM by conducting phase plane analysis and found that it represents each set of input values (V_1, \dots, V_N) as one unique and stable equilibrium point in its output space (R_1, \dots, R_N) when $\beta = 0$. Specifically, we solved the equilibrium state of each R unit by setting each differential equation (Eqs. 1-3) to zero, which results in the nullcline of each R unit as a function of the given activities of the other R unit, visualized in **Fig. 3D**. The nullclines of R_1 (solid) and R_2 (dashed) intersect at a unique point, regardless of equal or unequal input values (see different panels for examples of different inputs). This point indicates that the dynamical system, when receiving any positive inputs, can maintain an equilibrium where every unit maintains a steady level of activity. Linearization analysis around this point suggests that this point is attractive: given any initial values to the system, the activities of the units will converge into the unique equilibrium point (see **Methods Equilibria and stability analysis of the LDDM** for mathematical proof). The solution of the steady state of neural activity at the equilibrium (noted as R_i^*) reflects divisive normalization (Eq. 4), inheriting the property from the original DNM (LoFaro et al., 2014; Louie et al., 2014). The only difference at the equilibrium is the constant in the denominator ($B_G - \alpha$) introduced by baseline gain control and recurrent excitation; this change rescales the activity magnitudes but preserves normalized value coding.

$$R_i^* = \frac{V_i + B_R}{1 + B_G - \alpha + \sum_{j=1}^N \omega_{ij} R_j^*} \quad (4)$$

”

C. The performance of the three models (LDDM, DNM, RNM) on value normalization (Fig. 4) is inferred as LDDM and DNM do a good job at explaining the empirical data, but RNM does not. However, the fit results (as correlations) are quite good for all three, with the RNM's $r^2 = 0.892$. This is quite a good fit. Not sure that this is evidence that the RNM "does not capture well the magnitudes of neural activity as a function of V_{in} ."

We thank the Reviewer for bringing up this issue, as it touches on two important points we wish to emphasize. First, fitting the RNM to the value normalization data does a reasonably good job at replicating the patterns - but only at parameter values that take the RNM out of a WTA decision-making regime. This is a point that is emphasized in the Results text (p. 14, lines 289-292):

“Furthermore, fitting the RNM to the data results in a parameter regime that can no longer generate WTA competition; instead, the model predicts mean firing rates in a low-activity regime with a maximum value of 3.5 Hz (**Fig. 4-figure supplement 2**). These results suggest that RNM models cannot simultaneously support both normalized value coding and WTA selection regimes.”

Second, the Reviewer is correct that the RNM - when fit to the value normalization data - does a reasonable job of capturing the normalized coding. The difference between subtraction (inherent in the RNM) and division is mild, that's why the performance of the three models was not hugely different from each other. The qualitative difference between subtractive and divisive gain control is the curvatures predicted under lateral inhibition. The divisive models (LDDM and DNM) show concave decreasing curves when V_{out} increases; in contrast, the RNM shows linearly decreasing curves. The neural activity magnitudes also make a difference between the two types of inhibition across different levels of V_{in} . The divisive models (LDDM and DNM) predict slightly better for these fine-tuned patterns than the subtractive type of inhibition (in RNM).

In the revision, we have now tuned down our discussion of the advantage that the divisive models have and clarify more clearly the difference between divisive and subtractive forms. Please see the revised Results (p. 14, lines 284-288):

“We found that fitting the standard RNM with its standard four parameters (see **Methods**) cannot capture the pattern of neural activity as well as the LDDM and DNM (right panel in **Fig. 4B**) ($R^2 = .8920$). This small but clear difference in performance between model classes arises from the difference between divisive (DNM, LDDM) and subtractive (RNM) types of inhibition, with subtractive inhibition failing to capture the concave contextual effects predicted by divisive models.”

D. There appears to be an error in Figure 4B, the RNM plots. A vertical line of color dots, that appears to be a duplicate from the legend, looks to be cut and pasted into several parts of the graph (when the empirical data is supposed to be the same across all 3 panels in Fig. 4B).

We greatly appreciate the reviewer pointing out this mistake. There was a duplicate of the legend in **Fig. 4B**. We have now fixed the figure in the revised manuscript.

E. The subset plots in Fig. 5C are incredibly hard to read.

We have removed this subset in our revision since this information is available as a figure in a cited reference.

F. Figure 6B is not a Q-Q plot as stated in the caption (which measures empirical quintiles against theoretical quantiles as a means of assessing the type of underlying distribution of the data).

We thank the reviewer for noticing this error. They are correct, these plots are quantile probability plots (QPP), (Audley and Pike, 1965; Ratcliff and Tuerlinckx, 2002; Heitz, 2014). QPPs are particularly important in evaluating the fit of mathematical decision models because they show the distributional characteristics of the underlying RT and choice distributions.

We have corrected the description in the main text, methods, and legends for **Fig. 6** and **Fig. 6-figure supplement 1**.

G. The redundant use of red to represent both instantaneous change rates vectors and nullclines of R_2 in the same plots (e.g., Fig. 8B, E) makes the results really hard to see.

We thank the Reviewer for noticing this difficulty. We have now improved the visualization of the vector field (changed the color to black) to distinguish them from the nullclines in **Fig. 8B, E** as well as other figures with the same issue (**Figs. 8-figure supplement 1A-C** and **Figs. 8-figure supplement 2A-C**).

H. The end of the results introduces GABA manipulations in the model (beginning on line 510). However, as far as I read, there is no GABA-specific inhibition in the model. The differential equations that define the LDDM are simple inhibitory/excitatory units. There is no modeling of GABA dynamics themselves. So this shift in calling inhibition "GABA" is confusing.

We apologize for the confusion, and agree that the original shift to a GABA manipulation was a bit abrupt. In the revision, we have rephrased the GABA-related manipulation to be a more general inhibitory potentiation, mentioning a GABA agonist as an example of pharmacological manipulation. We have revised the related labels in **Fig. 10** and the Methods accordingly.

In Results (pp. 30-31), we have revised the text away from using GABA and towards a general potentiation of inhibition:

"Inhibitory potentiation distinguishes LDDM from earlier models

The architecture of disinhibition employed by the LDDM is more structured than the earlier non-selective inhibition used in most standard competition networks. This distinction gives rise to the novel prediction from the LDDM that the influence of global changes in inhibitory tone is non-selective during representation, but switches to input-selective after disinhibition is increased. This reflects a fundamentally novel prediction of this class of model. The LDDM contains two different types of inhibition, and thus its reaction to inhibitory potentiation depends on both the state of the disinhibitory network and the intensity of potentiation. To highlight the importance of that prediction, we implemented different levels of inhibitory connection weights in both the LDDM and the standard RNM.

At the neural level, the LDDM predicts a dissociable effect of potentiated inhibition on the primary (*R*) neuron's activity (**Fig. 10A**). During option representation (cue interval in fixed duration trials), potentiated inhibition increases both recurrent and lateral inhibition, leading to decreased firing rates and a weaker modulation by value in the *R* neurons. During option selection (go/choice intervals in fixed duration trials), local disinhibition increases WTA activity and decreases the late-stage representation of value. As an outcome, these changes produce a speeding up of RTs but a reduced choice accuracy (**Fig. 10B**). The expected differences between the control condition and the inhibitory potentiation condition would be evident in chronometric and psychometric curves across different levels of inputs effectively implementing a speed-accuracy tradeoff (**Fig. 10C**). Note that the qualitative predictions for inhibitory potentiation effects on RT and accuracy are robust to specific LDDM parameterizations (**Fig. 10D**). In contrast, in more traditional networks like the RNM that employ non-selective inhibition, potentiated inhibition suppresses the excitatory neural activities during the WTA competition (**Fig. 10E**). The suppression in neural coding in these models slows down RTs but does not affect choice accuracy (**Figs. 10F, G**), thus failing to replicate the observed speed-accuracy tradeoff. We note that these novel predictions that differentiate models which rely on structured disinhibition could be readily tested using modern optogenetic techniques."

I. The Methods introduce *E* and *I* signals as part of the LDDM, however, this is the first time these are mentioned (in fact the differential equations in the Methods differ with the inclusion of these signals, from the equations used in the Results). These seem key (and related to my prior point about GABA). Why are these excluded from prior descriptions of the model outside of the Methods?

We apologize for the confusion raised in the original Methods section, which arose from a lack of clarity in presentation. The *E* and *I* signals are part of an expanded model from which we derived the LDDM, which is the focus of this paper; we originally presented the expanded model at the start of the Methods to highlight its role in ruling out alternative model architectures. Thus, this part of the Methods was meant to explain how we selected the LDDM among other candidates. The results in the paper, and the rest of the Methods section, are entirely about the LDDM model. The expanded version of model testing thus describes a more general architecture, of which LDDM is a subset. We included two additional units, the "*E*" and "*I*", in addition to the units that remain in the current version of the LDDM. Exploratory analysis found that the redundant "*E*" and "*I*" units are not required to drive winner-take-all choice, thus we removed them from the circuit in the principle of parsimony.

We agree with the Reviewer that these details are minor in the scope of the paper, which centers around the LDDM, its architecture, and its capabilities. Thus, in the revised manuscript, we have reordered the Methods and moved the model exploration section to the end of the Methods in order to reduce confusion and distraction to the readers and kept the phase plane analysis at the beginning. In addition, we corrected multiple notation issues in this part, especially on the assumption of β , to align with the framework of the LDDM. For more detail in the changes regarding the notation of β and its function, please see the response to Reviewer #3.

Reviewer #2 (Public Review):

The aim of this article was to create a biologically plausible model of decision-making that can both represent a choice's value and reproduce winner-take-all ramping behavior that determines the choice, two fundamental components of value- based decision-making. Both of these aspects have been studied and

modeled independently but empirical studies have found that single neurons can switch between both of the aspects (i.e., from representing value to winner-take-all ramping behavior) in ways that are not well described by current biological plausible models of decision making.

The current article provides a thorough investigation of a new model (the local disinhibition decision model; LDDM) that has the goal of combining value representations and winner-takes-all ramping dynamics related to choice. Their model uses biologically plausible disinhibition to control the levels of inhibition in a local network of simulated neurons. Through a careful series of simulation experiments, they demonstrate that their network can first represent the value of different options, then switch to winner-takes-all ramping dynamics when a choice needs to be made. They further demonstrate that their single model reproduces key components of value-based and winner-takes-all dynamics found in both neural and behavioral data. They additionally conduct simulation studies to demonstrate that recurrent excitatory properties in their network produce value-persistence behavior that could be related to memory. They end by conducting a careful simulation study of the influence of GABA agonists that provide clear and testable predictions of their proposed role of inhibition in the neural processes that underlie decision-making. This last piece is especially important as it provides a clear set of predictions and experiments to help support or falsify their model.

There are overall many strengths to this paper. As the authors note, current network models do not explain both value- based and ramping-like decision-making properties. Their thorough simulation studies and their validation against empirical neural and behavioral data will be of strong interest to neuroscientists and psychologists interested in value- based decision-making. The simulations related to persistence and the GABA-agonist experiments they propose also provide very clear guidelines for future research that would help advance the field of decision-making research.

Although the methods and model were generally clear, there was a fair amount of emphasis on the role of recurrence in the LDDM, but very little evidence that recurrence was important or necessary for any of the empirical data examined. The authors do demonstrate the importance of recurrence in some of their simulation studies (particularly in their studies of persistence), but these would need to be compared against empirical data to be validated. Nevertheless, the model and thorough simulation investigations will likely help develop more precise theories of value-based decision-making.

We appreciate the Reviewer's thoughtful comments. These comments - especially about anatomic recurrence and its relationship to the parameter α - inspired us to think more about the uniqueness of the current circuit to others, especially the implications related to the parameters α (i.e., self-excitation) and β (i.e., local disinhibition). Recurrence is required to drive winner-take-all competition in the standard RNM of decision-making. However, we show here with both analytical and numerical approaches that recurrence helps WTA competition but is not necessary in our model. Instead, the key feature of the LDDM is to utilize disinhibition in conjunction with lateral inhibition to realize winner-take-all competition. That leads to many different predictions of the current model from the existing models, such as selective inhibition and flexible control of dynamics.

In response to the Reviewer's points and after careful consideration of the differential equations, we realized that in our model fitting, the α parameter fitting to zero does not necessarily mean recurrence should be zero. The α parameter shares a lot of similarity to the baseline gain control (parameter B_G in our revision), and thus is unidentifiable in the current dataset. In the interest of parsimony, we did not include the parameter B_G in the original manuscript, but now include it because it reveals the difficulty of interpreting fit α values as simply the level of recurrence.

Overall, disinhibition (β) in the LDDM is required for WTA activity while recurrence (α) can contribute but is not necessary; however, α is theoretically important for generating persistent activity, with the caveat that in the current framework there is an unclear relationship between fit α and recurrence. Regardless, we agree that the contribution of α to the LDDM framework is worth further testing and examining with future empirical data.

Please see response to recommendations below for a more detailed discussion of these issues.

Reviewer #2 (Recommendations for the authors):

The manuscript and figures are overall very clear but I have a few minor comments/questions mainly related to the role/need of recurrent connections in the R units:

1) In their simulations, the authors explore how the LDDM behaves under a wide range of different parameter values. However, there seems to be some correlation between the recurrent α and disinhibitory β behavior in situations where no persistent neural behavior is needed. These parameters have an influence primarily during the choice phases and both seem like they could account for winner-takes-all ramping-like behavior. The LDDM model fits in the data sets that only measure value-based and winner-takes-all phases suggesting that recurrence is not needed (α fits in both cases = 0; this may be different for the third case study of the Churchland et al. 2008 data set but I did not see any fit values). For value-based decision-making and WTA dynamics, is recurrence necessary?

The Reviewer makes an excellent point about the contribution of α and β in the LDDM circuit. Conceptually, they play similar roles during WTA selection: by enhancing the activity of an option-specific R neuron in an R activity-dependent manner (either through recurrent activation or disinhibition), both parameters contribute to the competition process. We have shown in our previous version of the manuscript that α and β have qualitative differences in predicting the dynamics of the circuit. Non-zero β is required for WTA competition; as the Reviewer intuitively, α will facilitate WTA competition but is not required (**Fig. 5E**). Since in the numerical simulation of WTA choice behavior, both α and β control the ramping-up speed of competition, they do show some level of collinearity in fitting the choice behavior (see the visualization of log-likelihood in the α - β space in revised **Fig. 6-figure supplement 1**).

However, spurred by the Reviewer's question, we conducted further analyses showing that α and β contribute differently to the shape of reaction-time (RT) distribution in the perceptual decision task. Increasing β decreases the skewness of the reaction time distribution; whereas, increasing α increases the skewness of the RT distribution (**Fig. R1**, below). Considering the capacity of the current article, we propose to systematically investigate the shape of RT distribution in our next paper, which we are currently drafting on the topic of speed and accuracy tradeoffs in decision-making. Oftentimes, when animals or humans prioritize speed in decision-making tasks, the shape of RT distribution appears as less skewed than in a condition prioritizing accuracy over speed. The speeding up of RTs and decreasing of skewness is consistent with increasing β instead of increasing α . We believe such evidence partially supports β instead of α being the parameter that flexibly changes over conditions and mental states. In contrast, α , which indicates the recurrence of the circuit and most likely an anatomical property of the microcircuit, should be relatively fixed.

While the points mentioned above suggest the importance of β in the model, we point out one important note of caution regarding interpretation of the fitted value of α . After considering the Reviewer's points about a different α between persistent activity (working memory) and valuation/WTA states, we re-examined the influence of incorporating a baseline or background input to the G neurons (B_G , as already exists for the R neurons as B_R). We have now included additional analyses in our revision which show that α is highly collinear with the baseline gain control level of the circuit B_G (please see the response below to Reviewer #3, point 3). Given this, α alone is generally unidentifiable under the current datasets; in fact, the term $(B_G - \alpha)$ is identifiable. Thus, the fitted value of α does not reflect the exact level of recurrence of the network. At this point we believe it is difficult to make any conclusions about how the level of recurrence changes across conditions or datasets and so temper our text in this area.

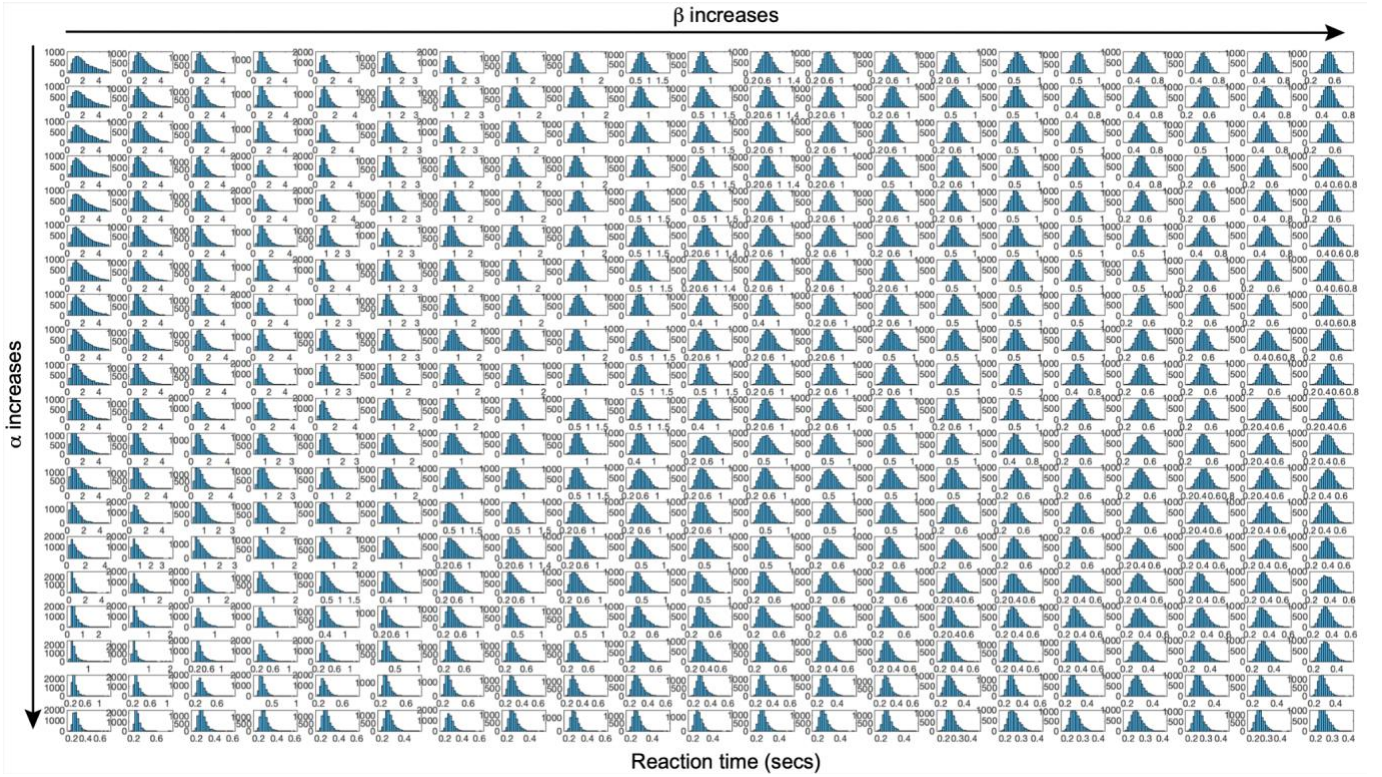


Figure R1. The shape of predicted reaction-time distribution over a wide range of α and β values by LDDM. Each grid indicates the predicted RT histogram normalized in the range of minimum and maximum RTs. The shape of RT distribution exhibits a pattern of increasing skewness when α increases and decreasing skewness when β increases.

2) Following the previous comment, can the model parameters be recovered in situations where there are only value- based and/or WTA dynamics (especially the alpha and beta parameters)? Showing that these parameters are recoverable would help support that these parameters have a specific and identifiable role within the network (see Wilson and Collins, 2019, eLife).

We agree with the Reviewer that parameter recovery is important to verify both the identifiability and specificity of model parameters. In this revision, we now present parameter recovery results for both the value normalization neural activity fit and the perceptual choice data fit. Both these parameter-recovery results suggest that the model parameters are generally specific and identifiable, with the exception noted below.

In terms of LDDM fits to the value normalization data, the refitted parameter $B_G - \alpha$ (see above) keeps consistent with the generated values within a reasonable range (from -20 to 20); and the refitted baseline input B_R is highly consistent with the generated values across a wide range (tested from 0 to 140). The best-fit parameters $B_G - \alpha = 3.82$ and $B_R = 71.53$. Both of them are in the range of high recovery. By shifting to $B_G - \alpha$ we thus maintain full identifiability as now noted in the manuscript

Model-recovery for the value normalization fits is presented in the revised manuscript in **Fig. 4-figure supplement 1** and briefly mentioned in the Results text (p. 13, lines 277-278):

“...parameter recovery analysis shows that the LDDM is highly robust in the data fitting, **Fig. 4-figure supplement 1**”

For the fits to perceptual choice data, we took a narrower approach since model fitting to every new (simulated) dataset takes ~3 days even on our high-performance cluster; conducting an analysis as in Wilson and Collins (2019) to examine, for example 100 pairs of simulated and fitted values would require an unreasonable amount of time (about one year). Fortunately, it is possible to examine how well the model can recover the set of best fitting parameters. To do so, we (1) generated a simulated dataset of reaction time and choice behavior using the best fitting parameters, and (2) re-fit the LDDM to that simulated dataset. We visualized the log-likelihood as the goodness of re-fit in the space of each pair of parameters across wide ranges of parameter values. All of the log-likelihood spaces show smooth and single-maximum topography (**Fig. 6-figure supplement 2**), similar to the likelihood spaces that we showed in fitting to the empirical data (**Fig. 6-figure supplement 1**). The parameter values on the peaks of log-likelihood, i.e., the best fitting parameters given the precisions of the grids, precisely matched the values on the peaks in fitting to the data. These results suggest that the model parameters are recoverable and identifiable. We would also like to point out that the well-featured model recovery may benefit from the quantile maximum likelihood estimation (QMLE) technique (Heathcote et al., 2002; Brown and Heathcote, 2003) we implemented here. QMLE captures the 9 quantiles of reaction-time distribution instead of the full distribution, and therefore reduces the influence of extreme values and outliers and prevents overfitting (Ratcliff, 1979; Ratcliff & McKoon, 2008).

Model-recovery in the revised manuscript is now presented in **Fig-6 supplement 2** and briefly mentioned in the Results text (p. 19, lines 408-410):

“A parameter recovery analysis indicated that the parameters are recoverable and identifiable within the network (**Fig. 6-figure supplement 2**).”

3) The authors demonstrate that recurrence can simulate persistence-like behavior in neural populations if $\alpha \geq 1$ in R units. The authors' fits found that $\alpha = 0$ in cases that did not require persistence (the value-based and WTA studies) - however, α would need to be ≥ 1 in cases that require persistent activity, suggesting that the network would need to modulate the levels of recurrence in R units when persistence is necessary. This does not seem plausible in my mind without some outside influence. How do the authors make sense of these differences?

The Reviewer makes a great point about how to interpret the α parameter and the plausibility of a changing level of recurrence. Given this point, for the revision we have carefully examined the recurrence parameter α with other parameters, particularly with a parameter corresponding to baseline or background input to G neurons (B_G), which is analogous to the baseline parameter for R neurons (B_R). In these new analyses, we find that α is highly collinear with the baseline gain control parameter B_G , in both fitting to the equilibrium activities during representation (revised Results, pp. 13-14) and predicting winner-take-all choice behavior (revised Results, pp. 18-19). In light of these results, the previous identification of α as zero in the original manuscript arose because we assumed $B_G = 0$. If we instead set B_G to a higher value (for example, an arbitrary value of 20), the best fitting α will increase and become larger than zero (see **Fig. R2** attached below). In the revision, we now show the collinearity test results between each pair of parameters, highlighting the collinearity between α and B_G (please see the revised **Fig. 6-figure supplement 3**). As mentioned above in the response to point (1), the key conclusion is that the fitted value of α does not reflect the exact level of recurrence of the network, and we believe it is difficult to make any conclusions about how the level of recurrence changes across conditions or datasets based on fit α values.

It is interesting to consider the source of the collinearity between α and B_G . We believe this collinearity may reflect E-I balance in the network, with α controlling the recurrent/excitatory degree of the network and B_G controlling the baseline/default status of inhibition to the network that balances the excitation. Given this collinearity, our analyses show that persistent activity in the LDDM requires a sufficiently large ($B_G - \alpha$) term (please see the revised analytical analysis for persistent activity; pp. 52-53). However, the collinearity between α and B_G in the model fittings means that the exact value of α or B_G cannot be identified given the existing datasets. Future empirical work

is needed to identify the features of these parameters. For example, one possible observation in the near future is to measure the neural activities of different types of neurons, e.g., primary neurons, SST/PV interneurons, and VIP neurons, using advanced neural imaging techniques (e.g., optogenetics, cell-specific targeting, high-density recording arrays, optical imaging). Since we propose that baseline gain control is linked to the activity of SST/PV interneurons, a direct test can be measuring the activities of SST and PV neurons across the full dynamics of decision-making tasks. That will help the identification of B_G , helping to dissociate its contribution from that of recurrence.

To address these issues, the revised manuscript includes the new analyses and results about the α and B_G collinearity, with the limitation on the identification of α and B_G discussed in the Results (p. 13, lines 278-282 and pp. 19-20, lines 419-422):

“Note that fitting to the current dataset is not able to differentiate the contributions of α and B_G to the neural dynamics (see proof in Methods); thus more empirical data will be needed to draw conclusions about the specific role of recurrent self-excitation in value coding. However, we do show below that self-excitation is critical for generating persistent activities (see section *Disinhibition controls point versus line attractor dynamics in persistent activity*).”

“Given the collinearity issue between B_G and α , the fitted value of α does not reflect the exact level of recurrence in the circuit, and future empirical data will be needed to differentiate how recurrence and baseline inhibition contribute to LDDM WTA selection.”

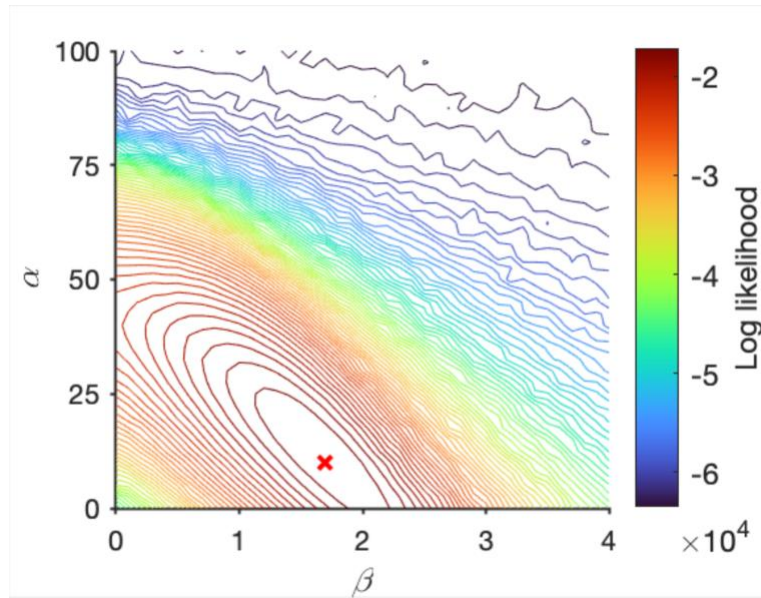


Figure R2. Fit α values depend on assumed B_G magnitude. The best fitting α value shifts higher when B_G was set to 20 instead of zero. The space shows the goodness of fit (log-likelihood) of LDDM to Roitman & Shadlen’s (2006) dataset as a function of α and β . Contour lines indicate the isolines of log-likelihood, with colors indicating value and the red cross indicating the maximal log-likelihood.

4) The manuscript and figures are overall very clear but I have one minor request: the caption in Figure 6 suggests that there is neurophysiological data being fit, but I see none in the figures (from what I can see, panels A-C are behavioral, and D-F are simulations). Could the neurophysiological data being fit be added to this figure (or more clearly identified if it is indeed there)?

We apologize for the lack of clarity. In **Fig. 6** (and supplements), the models were fit to behavior (choice and RT) only, and then models with the best fitting parameters used to generate the simulated neural dynamics. That is to say, the neural dynamics shown here do not result from a fit to the neural dynamics. This is now pointed out in the Results text (p. 20, lines 434-435):

“Notably, the LDDM – fit only to behavior – generates predictions about the underlying neural dynamics that can be compared to electrophysiological findings.”

In addition, we have revised the caption title in **Fig. 6** to convey this point more clearly (note that in response to other Reviewer points, the revised **Fig. 6** only includes LDDM results, while additional figure supplements present RNM and LCA model results). The relevant portion of the revised caption for **Fig. 6** states:

“**D.** The model with best-fitting parameters to the behavior replicates the neural dynamic features of the recorded neural activity.”

In regards to fitting to the neurophysiological data, we mentioned above that Adequately fitting the model(s) to neurophysiological data – particularly the fully dynamics of decision-related activity – will require more work and empirical observations, and given current technical limitations we do not address fitting the models to neural data in the current manuscript. However, we believe there are informative aspects of the approach we currently take – fitting the models to behavioral (choice and RT) data, and then examining the neural activity of the best fitting model units. We now quantify how well different model excitatory unit activity – with model parameters fit to behavior – matches empirically recorded data (from Roitman & Shadlen), focusing on how population average activity varies with stimulus level (coherence) at different timepoints in the trial. Please see our response to Reviewer #1, recommendation point 4 for more details.

Reviewer #3 (Public Review):

Shen et al. attempt to reconcile two distinct features of neural responses in frontoparietal areas during perceptual and value-guided decision-making into a single biologically realistic circuit model. First, previous work has demonstrated that value coding in the parietal cortex is relative (dependent on the value of all available choice options) and that this feature can be explained by divisive normalization, implemented using adaptive gain control in a recurrently connected circuit model (Louie et al, 2011). Second, a wealth of previous studies on perceptual decision-making (Gold & Shadlen 2007) have provided strong evidence that competitive winner-take-all dynamics implemented through recurrent dynamics characterized by mutual inhibition (Wang 2008) can account for categorical choice coding. The authors propose a circuit model whose key feature is the flexible gating of 'disinhibition', which captures both types of computation - divisive normalization and winner-take-all competition. The model is qualitatively able to explain the 'early' transients in parietal neural responses, which show signatures of divisive normalization indicating a relative value code, persistent activity during delay periods, and 'late' accumulation-to-bound type categorical responses prior to the report of choice/action onset.

The attempt to integrate these two sets of findings by a unified circuit model is certainly interesting and would be useful to those who seek a tighter link between biologically realistic recurrent neural network models and neural recordings. I also appreciate the effort undertaken by the authors in using analytical tools to gain an understanding of the underlying dynamical mechanism of the proposed model. However, I have two major concerns. First, the manuscript in its current form lacks sufficient clarity, specifically in

how some of the key parameters of the model are supposed to be interpreted (see point 1 below). Second, the authors overlook important previous work that is closely related to the ideas that are being presented in this paper (see point 2 below).

1) The behavior of the proposed model is critically dependent on a single parameter 'beta' whose value, the authors claim, controls the switch from value-coding to choice-coding. However, the precise definition/interpretation of 'beta' seems inconsistent in different parts of the text. I elaborate on this issue in sub-points (1a-b) below:

1a). For instance, in the equations of the main text (Equations 1-3), 'beta' is used to denote the coupling from the excitatory units (R) to the disinhibitory units (D) in Equations 1-3. However, in the main figures (Fig 2) and in the methods (Equation 5-8), 'beta' is instead used to refer to the coupling between the disinhibitory (D) and the inhibitory gain control units (G). Based on my reading of the text (and the predominant definition used by the authors themselves in the main figures and the methods), it seems that 'beta' should be the coupling between the D and G units.

1b). A more general and critical issue is the failure to clearly specify whether this coupling of D-G units (parameterized by 'beta') should be interpreted as a 'functional' one, or an 'anatomical' one. A straightforward interpretation of the model equations (Equations 5-8) suggests that 'beta' is the synaptic weight (anatomical coupling) between the D and G units/populations. However, significant portions of the text seem to indicate otherwise (i.e a 'functional' coupling). I elaborate on this in subpoints (i-iii) below:

(1b-i). One of the main claims of the paper is that the value of 'beta' is under 'external' top-down control (Figure 2 caption, lines 124-126). When 'beta' equals zero, the model is consistent with the previous DNM model (dynamic normalization, Louie et al 2011), but for moderate/large non-zero values of 'beta', the network exhibits WTA dynamics. If 'beta' is indeed the anatomical coupling between D and G (as suggested by the equations of the model), then, are we to interpret that the synaptic weight between D-G is changed by the top-down control signal within a trial? My understanding of the text suggests that this is not in fact the case. Instead, the authors seem to want to convey that top-down input "functionally" gates the activity of D units. When the top-down control signal is "off", the disinhibitory units (D) are "effectively absent" (i.e their activity is clamped at zero as in the schematic in Fig 2B), and therefore do not drive the G units. This would in turn be equivalent to there being no "anatomical coupling" between D and G. However when the top-down signal is "on", D units have non-zero activity (schematic in Fig 2B), and therefore drive the G units, ultimately resulting in WTA-like dynamics.

(1b-ii). Therefore, it seems like when the authors say that beta equals zero during the value coding phase they are almost certainly referring to a functional coupling from D to G, or else it would be inconsistent with their other claim that the proposed model flexibly reconfigures dynamics only through a single top-down input but without a change to the circuit architecture (reiterated in lines 398-399, 442-444, 544-546, 557-558, 579-590). However, such a 'functional' definition of 'beta' would seem inconsistent with how it should actually be interpreted based on the model equations, and also somewhat misleading considering the claim that the proposed network is a biologically realistic circuit model.

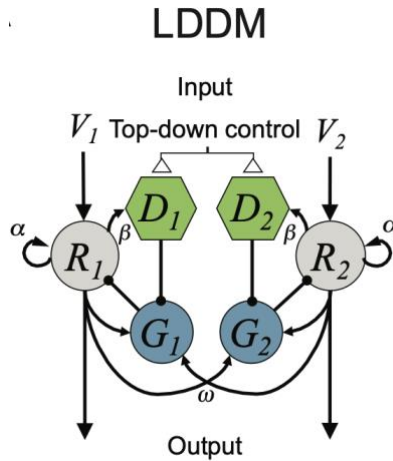
(1b-iii). The only way to reconcile the results with an 'anatomical' interpretation of 'beta' is if there is a way to clamp the values of the 'D' units to zero when the top-down control signal is 'off'. Considering that the D units also integrate feed-forward inputs from the excitatory R units (Fig 2, Equations 1-3 or 5-8), this can be achieved either via a non-linearity, or if the top-down control input multiplicatively gates the synapse (consistent with the argument made in lines 115-116 and 585-586 that this top-down control signal is 'neuromodulatory' in nature). Neither of these two scenarios seems to be consistent with the basic definition of the model (Equations 1-3), which therefore confirms my suspicion that the interpretation of 'beta' being used in the text is more consistent with a 'functional' coupling from D to G.

We thank the reviewer for pointing out this confusion. We apologize that the original illustrations (**Fig. 2A**) and the differential equations in Methods (Eqs. 5-8) did not convey very well our ideas. β is intended to reference the coupling from R to D , not a change in the weights between D and G units. We realize there was some confusion on this part due to inconsistency between our original figures, text, and supplementary material.

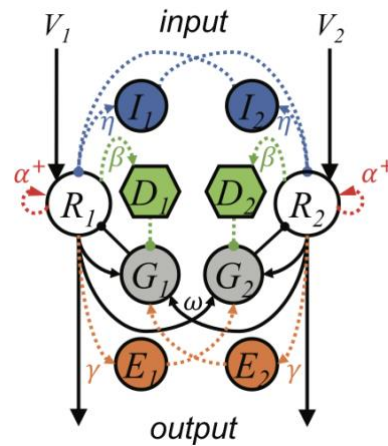
Given the lack of clarity in the previous version as well as the Reviewer's questions, we now emphasize that β represents a functional coupling between the R and D neurons. The biological assumption of the disinhibitory architecture is built based on recent findings that VIP neurons in the cortex always inhibit other neighboring inhibitory cells, such as SST and PV neurons, and consequently disinhibit the neighboring primary neurons (e.g., Fu et al., 2014; Karnani et al., 2014, 2016). We did not see evidence in the literature of fast-changing (anatomic) connections between VIP and SST/PV. However, there is evidence that the responsiveness of VIP neurons to excitatory neurons can be modulated by changing the concentrations of neuromodulators, such as acetylcholine and serotonin (Prönnke et al., 2020). While the stereotype of neuromodulator action is slow dynamics, recent findings show that for example basal forebrain cholinergic neurons respond to reward and punishment with surprising speed and precision (18 ± 3 ms) (Hangya et al., 2015) to modulate arousal, attention, and learning in the neocortex. Given the large number of studies that identify long-term projections and neuromodulatory inputs to VIP neurons (e.g., Pfeffer et al., 2013; Pi et al., 2013; Alitto & Dan, 2013; Tremblay et al., 2016), we believe that it will be more plausible to assume the connection weights between R and D in our case is quickly modulated within a trial.

To clarify this issue in the revised manuscript, we made the following corrections:

1. We repositioned the β parameter in **Fig. 2A** between the connection from R to D , to align the description of β modulating R to D in the main text.
2. We modified the differential equations 5-8 (now numbered as Eqs. 28-32) in Methods (pp. 61) to include the disinhibitory unit D as an independent control from the inhibitory unit I , in order to be consistent with the disinhibitory D units in LDDM. Such a change makes tiny differences in the model predictions (please see dynamics simulated after the change in **Fig. 2-figure supplement 1B**).
3. We updated the neural circuit motif in **Fig. 2-figure supplement 1A** accordingly.



The revised Figure 2A



The revised Figure 2 – figure supplement 1A

2) The main contribution of the manuscript is to integrate the characteristics of the dynamic normalization model (Louie et al, 2011) and the winner-take-all behavior of recurrent circuit models that employ mutual inhibition (Wang, 2008), into a circuit motif that can flexibly switch between these two computations. The main ingredient for achieving this seems to be the dynamical 'gating' of the disinhibition, which produces a switch in the dynamics, from point-attractor-like 'stable' dynamics during value coding to saddle-point-like 'unstable' dynamics during categorical choice coding. While the specific use of disinhibition to switch between these two computations is new, the authors fail to cite previous work that has explored similar ideas that are closely related to the results being presented in their study. It would be very useful if the authors can elaborate on the relationship between their work and some of these previous studies. I elaborate on this point in (a-b) below:

2a) While the authors may be correct in claiming that RNM models based on mutual inhibition are incapable of relative value coding, it has already been shown previously that RNM models characterized by mutual inhibition can be flexibly reconfigured to produce dynamical regimes other than those that just support WTA competition (Machens, Romo & Brody, 2005). Similar to the behavior of the proposed model (Fig 9), the model by Machens and colleagues can flexibly switch between point-attractor dynamics (during stimulus encoding), line-attractor dynamics (during working memory), and saddle-point dynamics (during categorical choice) depending on the task epoch. It achieves this via a flexible reconfiguration of the external inputs to the RNM. Therefore, the authors should acknowledge that the mechanism they propose may just be one of many potential ways in which a single circuit motif is reconfigured to produce different task dynamics. This also brings into question their claim that the type of persistent activity produced by the model is "novel", which I don't believe it is (see Machens et al 2005 for the same line-attractor-based mechanism for working memory)

We thank the Reviewer for pointing out the conceptual similarities between the LDDM and the Machens Romo Brody model, and now include a discussion of the link between the two early in the revised Discussion (p. 38, lines 826-837). Please see response to recommendations below for a more detailed discussion of this point.

2b) The authors also fail to cite or describe their work in relation to previous work that has used disinhibition-based circuit motifs to achieve all 3 proposed functions of their model - (i) divisive normalization (Litwin-Kumar et al, 2016), (ii) flexible gating/decision making (Yang et al, 2016), and working memory maintenance (Kim & Sejnowski, 2021)

The Reviewer notes several relevant papers, and we have now discussed them and their relationship to the LDDM in a revised Discussion section (pp. 35-36). Please see response to recommendations below for a more details.

Reviewer #3 (Recommendations for the authors):

1) I think the authors will have to rewrite parts of the manuscript to address the major concerns I raise in the public review - (i) especially to clarify the precise interpretation of the single key parameter that determines the behavior of the model, and (ii) point out the connection to previous work (Machens et al 2005, Yang et al, 2016, Litwin-Kumar et al 2016) emphasizing the specific ways in which their work is an advance on these previous studies.

We thank the Reviewer for these suggestions, and have significantly rewritten parts of the manuscript to address these concerns.

One major point was the conceptual interpretation of the β parameter, which in the LDDM controls the degree of disinhibitory drive in the network and consequently governs the shift between value representation and WTA selection, as explained in the response to Public Review above.

The Reviewer rightly points out connections to past literature, both in terms of the Machens et al mutual inhibition model and its flexible dynamic states and other existing models with disinhibitory motifs that share some functions with the LDDM. Regarding the Machens model, it is an important related model: like the LDDM, it models a decision process with mutual inhibition as the key competitive mechanism, and it can flexibly reconfigure its dynamics to transition from stimulus encoding (point attractor) to working memory (line attractor) to WTA selection (saddle point). We do note that there are some key differences between the LDDM and the Machens model: (1) it models a sequential two-interval decision rather than a simultaneous decision, (2) while it achieves a reconfiguration of state via changes in external input, as in the LDDM's top down activation of disinhibition, it requires a distinct switch in functional connectivity between inputs and decision circuit elements, and (3) disinhibition plays an ancillary rather than central role (it is presented as a possible input-switching mechanism). We have added new text about the Machens et al model in the Discussion section (p. 38, lines 826-837):

“The LDDM achieves the flexible reconfiguration of line attractor and point attractor states under the control of disinhibition, suggesting that attractor dynamics might not be a fixed property of a network; rather, it may be adaptive and controllable by a top-down signal operating via gated disinhibition. Of course, similar reconfiguration has been achieved by other important circuit mechanisms that have been well-described. For example, a mutual inhibition network can capture the different regimes of sequential two-interval decision-making – stimulus loading, working memory, and comparison – by assuming a flexible reconfiguration of the external inputs (Machens et al., 2005). Similar to the LDDM, this model can transition between point attractor (initial stimulus encoding), line attractor (working memory), and saddle point (comparison) dynamics. Interestingly, disinhibition may also play a role in this model, by providing a theoretical mechanism to switch the routing of external inputs within the circuit, which drives the switch from line attractor to comparison dynamics.”

As the Reviewer notes, there are existing models that show a link between disinhibition and functions integrated together in the LDDM (divisive normalization, working memory, and gating/selection). We agree that it is important to relate the LDDM to prior work, and have added new text to the Discussion that discusses the relevant models raised by the reviewer (Litwin-Kumar et al, Kim and Sejnowski, Yang et al) as well as other predecessor literature (pp. 35-36, lines 761-779):

“The contribution of LDDM relative to existing disinhibition models

Disinhibition has been previously linked in separate models to several of the computational functions that are exhibited in a unified manner by the LDDM. For example, a computational model employing dendritic disinhibition captures flexible information routing in a context-dependent decision task, with dendritic disinhibition gating on specific inputs to a circuit while gating off other pathways (Yang et al., 2016). However, disinhibition plays a different role in this model (context-dependent input gating) from that employed in the LDDM (transition from value coding to WTA selection and mutual competition). In another example, PV neuron activation within a disinhibitory circuit motif can produce a divisive normalization of tuning curves in a model of visual cortex (Litwin-Kumar et al., 2016). This specific model of division, however, arises from different circuit mechanisms than those we employ, such as reduced tuned input and firing rate nonlinearities. Finally, disinhibition has also been proposed to underlie the long timescales of information processing seen in working memory, as enhancing inhibitory-to-inhibitory connections stabilize temporal dynamics and improve working memory performance in recurrent neural networks (R. Kim & Sejnowski, 2021). One other notable difference between previous research and our current work is that disinhibition in past models typically contributes to a specific function (e.g., input gating, categorical selection, working memory, etc.), whereas disinhibition in the LDDM both mediates a transition from value coding to WTA selection and plays an integral role in the selection process itself. Taken together, previous

results and our current work reinforce the importance of incorporating disinhibition in circuit models of decision-making.”

2) It would also greatly help if the usage of notation is made consistent throughout the paper. For instance, in the figures and the equations in the main text (Equations 1-3), disinhibition is denoted as D , but in the methods (Equations 5-8) and the supplementary figures (Figure 2, Supplementary Figure 1) it is denoted as I .

We thank the Reviewer for pointing this issue out. This confusion arose in the original manuscript because the equations in the Methods referred to a more general expanded version of the model, from which we derived the circuit model presented in the paper (LDDM). As we note in the reply to the Public Review above, we have corrected the Equations 5-8 to become Equations 28-32, to convey our idea more clearly. In the expanded models, we added D units to the testing motifs, separated from the I units in the old version. D units represent disinhibition modules for local disinhibition and I units represent inhibitory modules for cross inhibition. In the revised version, the meaning of D units in the expanded models is now consistent with the meaning of D units in the LDDM, throughout the Results and the Methods sections. Please see the revised Methods (pp. 60-62).

3) I appreciate that the authors also studied a more general and 'extended' version of their model (of which the LDDM is a special case) and explore how it behaves in different regions of parameter space (Figure 2, Supplementary Figure 1). However, I found the general description of their extended model quite confusing, particularly, some of the design choices. For instance, the extended model consists of additional excitatory units (E) that are referred to as 'gain control boost loops'. These are never mentioned in the main text and their purpose for the overall story of the paper seems somewhat unclear to me. Since the R units already have projections to both 'local' and 'lateral' gain units (through ' ω ', Fig 2A), couldn't the E units simply be replaced by stronger self-recurrence on the R units?

We apologize for any confusion in the original version of the manuscript, which likely arose because the previous version overemphasized the role of the extended model. Our paper focuses on the LDDM: a simple lateral inhibition model with recurrent excitation and an activatable within-option disinhibition. However, to emphasize that this model architecture was not selected in an arbitrary fashion, we included a brief synopsis of how we narrowed down a broader range of possible models to the version characterized in the rest of the paper (LDDM), based on desired functional characteristics (e.g., WTA activity).

The presentation of the extended version of the model is intended to show all possible modifications that were tested, and eliminated, before we settled on the LDDM architecture. In the main text, we do not discuss other motifs in the service of the readability of the paper. We realize that there were also some problems with notation inconsistency between the original Methods section and the original Results section. We have now corrected these notation problems, as we mentioned in the above point and in our response to the Public Review.

Regarding the E units, the reviewer is correct: adding an E unit to project selectively to the lateral G units will be very similar to selectively changing the connection weights matrix ω_{ij} . However, it will be different from simply changing the self-excitation on the R units (i.e., α), since each R unit receives gain control from lateral R units (mediated via G units) but also receives gain control from itself (via the local projection to G). As an outcome, changing the strength of self-excitation is not alone sufficient to break the balance between the two R units, i.e., will not lead to winner-take-all competition. Only when the ω_{ij} matrix is asymmetric - equivalent to introducing a new E unit to selectively target the lateral G - will the circuit be able to break the balance of self gain control and lead to winner-take-all competition.

We have now clarified these issues in the revised Methods (pp. 60-62):

We tested a series of motifs and found local disinhibition is critical for the integration of normalized valuation and choice functions. To do this, we tested four types of modifications that might enhance mutual competition between the option-specific local sub-circuits (**Fig. 2-figure supplement 1A**): a) *Recurrent self-excitation* (loops weighted by α), with self-amplification of each R unit, a property shown to be important for mutual competition in the RNM. b) *Local disinhibition* (loops weighted by β), which is the focus of the main text, mediated through disinhibitory units (D); the function of a D unit is to inhibit the gain control G unit in the local sub-circuit therefore release inhibition on the local R units. c) *Cross inhibition* (loops weighted by η), which directly inhibits the lateral R units through inhibitory units (I) to implement mutual inhibition. d) *Lateral gain control boost* (loops weighted by γ), which is mediated through excitatory units (E) to boosts the lateral G , therefore drives higher gain control on the lateral R than the local R (i.e., asymmetric gain control) and realizes mutual inhibition.

To see which type of modification(s) is/are critical for integrated value normalization and choice, we tested different combinations of these modifications on the original DNM circuit. The full model with all modifications can be described by a set of differential equations (Eqs. 28-32):

$$\tau_R \frac{dR_i}{dt} = -R_i + \frac{V_i + \alpha R_i - I_i}{1 + G_i} \quad (28)$$

$$\tau_G \frac{dG_i}{dt} = -G_i + \omega \sum_{j=1}^N R_j + E_i - D_i \quad (29)$$

$$\tau_D \frac{dD_i}{dt} = -D_i + \beta R_i \quad (30)$$

$$\tau_I \frac{dI_i}{dt} = -I_i + \eta R_i \quad (31)$$

$$\tau_E \frac{dE_i}{dt} = -E_i + \gamma R_i \quad (32)$$

where $i = 1, \dots, N$ designates choice alternatives, each of which receive input V_i , and τ_R , τ_G , τ_D , τ_I , and τ_E are the time constants for the R , G , D , I , and E units. The weights ω represent the coupling strength between excitatory units R and gain control units G , the parameters α , β , η , and γ control the active state of recurrent excitation, local disinhibition, cross inhibition, and lateral gain control boost loops, respectively.

The active and inactive states of the four types of loops can be combined into $2^4 = 16$ possible models. Example dynamics were shown in **Fig. 2-figure supplement 1B** for each type of model. When local disinhibition (β) is off (left two columns), the model generates WTA dynamics only when cross inhibition (η) is on. But the maximum activity in the late stage is still restricted to a value lower than the phasic peak during the early stage, contradicting empirical findings that the late stage decision threshold is usually higher than activity in the early phasic peak (Churchland et al., 2008; Kiani et al., 2008; Kiani & Shadlen, 2009; Louie et al., 2011; Roitman & Shadlen, 2002; Rorie et al., 2010; Shadlen & Newsome, 2001; Sugrue et al., 2004). This restriction arises because, with only cross inhibition, local option gain control is not released; this release requires local disinhibition. With local disinhibition on ($\beta > 0$, the right two columns), the models generate WTA dynamics with high activity in the late stage to reach the decision threshold. This is robust even without any other modifications (see the panel with η and γ off), highlighting the role of local disinhibition in generating WTA competition. For the sake of simplicity, we omitted other non-essential modifications and kept only the loop of local disinhibition. Because recurrent excitation is important for persistent activity and exists widely in cortical circuits, we retained it as well. The modified DNM model with local disinhibition and recurrent self-excitation is the primary model (LDDM) characterized in the current work.”

4) The most interesting analyses in the paper are where the authors fit the circuit model to neurophysiological data. The authors then report the values of the fitted parameters and also perform the

model comparisons by reporting AIC/likelihood ratios. However, if possible, it would be very informative to also visualize the optimization surface of these fits to understand whether some of the free parameters trade-offs against one another, as I think that would affect the overall conclusions drawn in the paper, and also convince me about the robustness of the fitted parameters.

We agree with the Reviewer that examining likelihood surfaces is informative for understanding the relationship between parameters in fitting the dataset at hand (note that we fit the model to the behavioral data, with neurophysiological activity derived from the behaviorally-fit model). In the revision, we now visualize the log-likelihood in the spaces of different pairs of parameters as shown in new figure supplements (**Fig. 6-figure supplement 1**). All of the spaces show smooth and single maximum topography, which suggests the model fitting is robust and stable. All of the parameters show no extreme collinearity and identifiable maximum value. α and β show mild collinearity, we included a discussion of this point in the response to Reviewer #2, recommendation point 1 and in **Fig. R1**. The best fitting parameters visualized in the grids precisely match the best-fitting parameters given the precision of the grids.

5) A somewhat more open-ended question is about the choice of the time constants for the 3 types of units in the model (R, G, and D), which appear to be fixed to a value of 100ms for all the results presented in the manuscript. Can the authors justify this choice? Considering that SST (which, I presume are the gain control units G) and VIP neurons have fundamentally different conductance profiles and are known to show an entire range of spiking patterns (Tremblay et al, 2016), is it justifiable to assume that their time constants all have the same value?

We agreed with the Reviewer that the time constants for different units could (or should) possibly be different, given the biological assumptions of different neuronal types. Thus, we fitted the τ of each unit as free parameters in the model fitting to Roitman & Shadlen's data. We used fixed τ (100 ms) only for visualizing example dynamics when they are not quantitative results (e.g., the example dynamics shown in **Fig. 2**). In addition, in the equilibrium analyses, time constants do not affect the equilibria of the circuit and thus the exact values we assumed on the τ does not impact our conclusions. We have clarified this issue in our revised Results (p. 19, lines 402-407) and Methods (p. 44, lines 945-946).

"The model is then reduced to seven parameters: recurrent excitation weight α , local disinhibition weight β , noise parameter σ , input value scaling parameter S , and time constants τ_R , τ_G , and τ_D (see Methods for model-fitting details). Predictions of the best fitting model are shown in **Fig. 6A** (best fitting parameters: $\alpha = 0$, $\beta = 1.434$, $\sigma = 25.36$, $S = 3251$, $\tau_R = .1853$, $\tau_G = .2244$, and $\tau_D = .3231$)."

" τ_R , τ_G , and τ_D were set as the same value of 100 ms only for non-quantitative visualization purposes and fitted independently as free parameters in the model fittings"

6) In general, the presentation of the figures can be improved:

a) In Figure 2-figure supplement 1, I should be replaced by D. Also, the parameter gamma seems to be missing from the rows in subpanel B of this figure.

At the Reviewer's suggestion, we have improved these illustrations and replaced the missing information. Please see the revised **Figure 2-figure supplement 1**. Thanks for pointing this out.

b) In Fig 5, it's hard to follow which subpanels are the 'main' subpanels and which ones are the insets.

We have revised **Fig. 5** on the sake of clarity, removing the inset subpanels since these are known results/properties from the RNM and have been reported in the cited literature.

c) The legend of Fig 4B (right column) seems to have an extra set of dots (ones that indicate the legend of V_{in})

We have corrected the mistake in the illustration, thank you.