nature portfolio

Peer Review File

Computational and neural mechanisms underlying the influence of action affordances on value learning

Corresponding Author: Dr Sanghyun Yi

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Yi and O'Doherty present a computational, behavioral, and neuroimaging-based investigation of the role of action affordances in guiding action selection during value-based decision making. They use behavioral and brain responses to arbitrate between three primary computational models. The results suggest that valued-based and affordance-based action selection systems are separate and competitive, with a third meta-controller arbitrating between them on a dynamic basis.

I found this to be a remarkably strong paper on multiple fronts. The computational modeling formalizes a clear set of alternative explanations of the psychological phenomenon of interest. The combination of multiple measures – reaction times, choices, and brain responses – provides strong convergent evidence. The experimental method is creative and naturalistic, yet well-controlled. And the overall research question is both important and understudied. Thus, I think that this work will make a very valuable contribution to the literature. I have a few suggestions, below, for how it might be further enhanced:

1. The paper places a strong emphasis on the visuomotor nature of action affordances, but actions can be construed at multiple levels of abstraction. To take the authors' own example, a computer keyboard does indeed afford pressing or poking actions. However, it also affords higher-level actions, such as writing, which are not dependent on the execution of a particular motor program (i.e., one can write through typing or via a pen or electronic stylus, or even via dictation). In everyday life, it is often these higher level affordances which guide decision making, because they are more directly tied to the goals that people are attempting to achieve. It might be useful to acknowledge this as outside the scope of the present investigation, yet still worthy of future investigation.

2. Even basic motor affordances must often be learned. Again returning to the keyboard example, the knowledge that a keyboard affords "pressing" is tied to experience with that object – for example, knowing that the keys are spring-loaded. The same is true for many of the objects investigated in the present study. In future work, it might be helpful to use cross-cultural differences in affordance knowledge to distinguish between more fine-grained computational models. For example, an experienced user of chopsticks likely perceives far more affordances in them than someone who has never used them – and people of both types are likely available for recruitment in the authors' environs. This will allow future studies to disambiguate whether the affordances under study are habit like or driven by more innate processes.

3. Many objects offer not just action affordances, but also affective affordances. For example, a can of beer offers not just the motor affordances of grasping or drinking, but also the affective affordance of drunkenness. The effects of such affective affordances – although rarely termed as such – has been extensively investigated in the substance use, addiction, and nutrition literatures within psychology. It would be helpful to make a connection between the present findings and those literatures, because I believe that they could benefit significantly from the conceptual clarity and computational specificity of the current approach (and this paper may receive more – deserving – attention as a result).

Reviewer #2

(Remarks to the Author)

The study by Yi and O'Doherty investigates the computational and neural mechanisms underlying the influence of action affordance on value learning. By combining functional neuroimaging, psychophysics, and computational modeling, they aim to understand how action affordance shapes value-based learning in novel environments. Their main finding suggests that action affordance operates as an independent system that concurrently guides action selection alongside value-based decision-making. These two systems engage in a competitive process to determine the best action. Notably, the authors demonstrate that different brain regions contribute to this process: the pre-SMA encodes the performance difference between the action-affordance and value-based decision-making systems, while the posterior parietal cortex (PPC) integrates predictions from both systems to facilitate action selection. Overall, this study provides important and novel insights into how the brain integrates information from multiple sources to optimize action selection. It fits within the scope of Nature Communications, but there are some points that the authors can address to further improve their study.

1. The authors describe how the brain integrates information about the affordance and the value of action. There are a number of previous studies that explored how the brain integrates information from multiple sources to make decisions. For instance, there are studies by Paul Cisek, Richard Andersen, Paul Glimcher and others (see below), which showed that the brain integrates dynamically information from multiple sources (such as value and effort/cost) when making decisions. Is it likely that the findings of the Yi and O'Doherty study reflects the integration of expected reward and effort cost to perform an action? This will be also consistent with the concept of desirability value, according to which the posterior parietal cortex integrates information from disparate sources, such as expected reward and effort cost. I suggest that the authors describe the previous studies by Cisek, Andersen, Glimcher, Schrater and others and explain how different are their findings from previous studies that have already reported that the brain integrates information about the affordance of action (i.e., cost/effort) and the value associated with that action.

• Canaveral CA, Lata W, Green AM, Cisek P. Biomechanical costs influence decisions made during ongoing actions. J Neurophysiol. 2024 Jul 11. doi: 10.1152/jn.00090.2024. Epub ahead of print. PMID: 38988286.
• Cos I, Duque J, Cisek P. Rapid prediction of biomechanical costs during action decisions. J Neurophysiol. 2014 Sep 15;112(6):1256-66. doi: 10.1152/jn.00147.2014. Epub 2014 Jun 3. PMID: 24899673.
• Christopoulos V, Bonaiuto J, Andersen RA. A biologically plausible computational theory for value integration and action selection in decisions with competing alternatives. PLoS Comput Biol. 2015 Mar 24;11(3):e1004104. doi: 10.1371/journal.pcbi.1004104. PMID: 25803729; PMCID: PMC4372613.
• Christopoulos V, Schrater PR. Dynamic Integration of Value Information into a Common Probability Currency as a Theory for Flexible Decision Making. PLoS Comput Biol. 2015 Sep 22;11(9):e1004402. doi: 10.1371/journal.pcbi.1004402. PMID: 26394299; PMCID: PMC4578920.
• Kable J, Glimcher P (2009) The neurobiology of decision: consensus and controversy. Neuron 63: 733–745. pmid:19778504

2. It will be also interested in assessing if there is any temporal difference between the two systems. Does the brain first assess the affordance of the action before the value or the other way? fMRI has a low temporal resolution, so it might not be possible to look temporal differences between the two systems, but it will be interested in trying to assess it.

3. Can the author assess whether and how the brain encodes the competition between available actions? For instance, it would be interesting to explore differences in brain activation between quick and slow responses. Previous studies have shown that reaction time reflects the level of competition between actions; that is, the longer it takes to respond, the stronger the competition between the available actions.

Minor points.

1. Please add line numbers in the manuscript. It is easier to review it.
2. The authors need to improve the quality of writing in some parts of the manuscript. For instance, there are some parts that there is no space between the sentences (e.g., last line in page 6, "…experiment.Because…").

Reviewer #3

(Remarks to the Author)
This study must have been a massively laborious undertaking, combining RL, mode-based fMRI, and automated decoding of gesture-based response. It is thus very unfortunate that I do not find the research question noteworthy. The interaction between action affordability and value-based decision-making is a very interesting topic, but the experimental paradigm in this study is incredibly artificial, at least for two reasons:

1. It is unclear if this task, in which participants see an object associated with (for instance) a pinching gesture and instead make a clenching gesture to earn a probabilistic reward, captures any real-world learning problems.

2. The experiment was divided into blocks, in each of which participants learned the most rewarding actions for four items, each corresponding to one of the four conditions (CH, CL, IH, and IL). The descriptions of instructions given to participants

are entirely missing in this manuscript, but even if they were not told anything about reward contingencies, it is natural for them to eventually start trying to figure out which items are in the congruent or incongruent condition in each block (it is perhaps one of the most rational strategies they can take).

It is unclear the degree to which the reported findings are generalizable beyond this experimental paradigm. The point 2 is particularly crucial for the significance of the modeling results; the "performance-based arbitration model" is effectively equivalent to first testing if the most affordable action is the most rewarding one (which is the case 50% of the time) and trying the other two actions if not. The artificiality of the experimental paradigm heavily limits my enthusiasm for this manuscript. It does not help that the authors did not conduct much analysis to demonstrate that their complicated models characterize the behavioral data better than simpler, perhaps heuristic, accounts (such as the one I described above).

Below are some of the specific issues I had with the modeling approach:

3. The models seem overparameterized. Just an example -- even in the simplest RL model, the authors had b_pinch, b_clench, and b_poke as free parameters. They cannot be uniquely estimated. Forgive my frankness, but I am surprised that the authors did not catch this, and such a mistake casts doubts on the entire modeling results. One suggestion would be to make the code publicly available at this point for transparency.

4. The authors make a lot of specific parametric assumptions in modeling, for example how affordance compatibility is linearly scored based on the post-task survey data. It is unclear to what degree the model comparison results depend on these (mis)specifications.

5. The authors examined an alternative model, in which learning rates differ between congruent and incongruent conditions. They rejected it solely based on the lack of population-level difference in learning rate estimates, but it would make much more sense to reject it based on its performance (e.g. log evidence).

6. The authors wrote: "We didn’t use the hierarchical Bayesian inference function of the toolkit because model simulations using the hierarchically estimated parameters couldn’t reproduce the characteristics of the real data" (p.24). This sentence is vague (what characteristics?), but I am not convinced that it is a scientifically sound and unbiased reason not to use the hierarchical model, particularly given the common observation in the literature that it tends to help estimate parameters more reliably.

7. In fMRI analysis, the wrote wrote: "we found positive correlations between each participant’s propensity to choose the most rewarding actions and the extent to which BOLD activity in the identified arbitration regions can be explained by our model-derived arbitration variables" (p. 15). Doesn't it simply mean that, for those whose behavioral data was well characterized by a specific model, their MRI data was also well characterized by the same model?

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have thoroughly addressed all the key points raised in my review, and I appreciate their thoughtful revisions. I have no further concerns and recommend the manuscript for publication.

Reviewer #4

(Remarks to the Author)
In this paper the authors explore how reward information and action affordances co-shape value-based learning.
This is an interesting research question, which also has applied implications.
The reported experiments are well designed and reported; and the same holds for most analyses in the manuscript.

My main concern is whether the authors have sufficient evidence to back up their main computational claim. Indeed, the behavioural data rule-out computational models in which the action affordance bias is fixed over time. The authors proceed and propose a series of models in which a dynamic bias emerges from cognitive control mechanisms. However, can the authors rule out the possibility that the affordance bias has its own (passive/ decaying) temporal dynamics? Incorporating a time-varying bias could potentially "patch-up" the simpler models the authors tried out. Where is the evidence that there is an active control mechanism causing this time-varying bias? Perhaps there are certain qualitative aspects of the behavioural data that would lend support to the arbitration models over a time-varying bias. In some, there seems to be another class of models in-between the fixed and arbitration ones that the authors need to assess.

Other comments:

-- The modelling endeavour is commendable, however the exposition of the modelling results can be better streamlined. For instance, comparing a Bayesian and RL model early on but then sticking with an RL formulation for the arbitration models is not well motivated/ explained.

-- It is not clear to me that the Affordance Prediction Error is a meaningful quantity. It combines incommensurable quantities: affordance compatibility and reward value. I also think that the notion of reliability is relatively fixed for affordance reliability. Reward reliability does increase as a function of learning, so I was wondering whether the authors have tried a simpler reliability model which resorts to the affordance bias only when value-based reliability is low (at the early trials).

-- The authors state that the RT behavioural effects are indistinguishable in early and late trials. However, the choice effects manifest differently at early trials (larger affordance bias) relative to late trials. It is not clear whether these two statements are at odds with each other.

-- At the end of section 2.3 the authors acknowledge that their experimental paradigm does not allow them to delineate the temporal dynamics of processing value information and action affordances. I agree with their statements. Across these lines, I find that the slow vs. fast RT analyses is not informative. This analysis could be informative if we were sure that all slow responses are attributed to a higher decision boundary (with reference to an evidence accumulation framework). However, slow responses can also be attributed to other factors (low arousal levels or drift rate, longer non-decision time etc) which limits the inferential potential of a median split analysis in this paradigm.

Reviewer #5

(Remarks to the Author)
This is an impressive multi-modal approach to investigating the interaction of action affordability and value-based decision making in a methodologically and theoretically strong paper. The authors admirably integrated the helpful comments by the reviewers in this revision, which has clarified several points about the strength of the modeling approach, the significance of the question, clarification on the methods and generalization of the paradigm/results. Helpful additions include supplementary figures 17-19, testing of new models (e.g., win-stay-lose-shift strategy) and addressing concerns about overparameterization. The authors also made their codebase publicly available to enhance transparency. Overall, this revised paper will make a significant contribution to the literature and is well-suited for publication in Nature Communication.

Version 2:

Reviewer comments:

Reviewer #4

(Remarks to the Author)
The authors engaged in-depth with my comments and addressed all my concerns. I have no further concerns.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Yi and O'Doherty present a computational, behavioral, and neuroimaging-based investigation of the role of action affordances in guiding action selection during value-based decision making. They use behavioral and brain responses to arbitrate between three primary computational models. The results suggest that valued-based and affordance-based action selection systems are separate and competitive, with a third meta-controller arbitrating between them on a dynamic basis.

I found this to be a remarkably strong paper on multiple fronts. The computational modeling formalizes a clear set of alternative explanations of the psychological phenomenon of interest. The combination of multiple measures – reaction times, choices, and brain responses – provides strong convergent evidence. The experimental method is creative and naturalistic, yet well-controlled. And the overall research question is both important and understudied. Thus, I think that this work will make a very valuable contribution to the literature. I have a few
suggestions, below, for how it might be further enhanced:

We would like to sincerely thank the reviewer for their thoughtful feedback. We appreciate your positive remarks on our work.

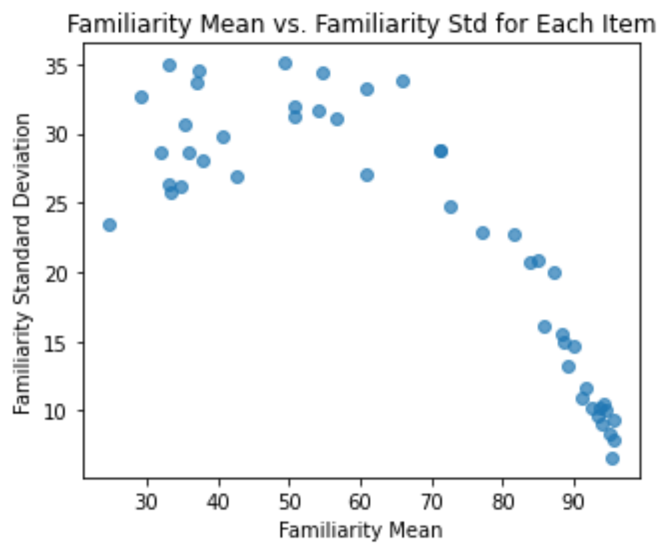
1. The paper places a strong emphasis on the visuomotor nature of action affordances, but actions can be construed at multiple levels of abstraction. To take the authors' own example, a computer keyboard does indeed afford pressing or poking actions. However, it also affords higher-level actions, such as writing, which are not dependent on the execution of a particular motor program (i.e., one can write through typing or via a pen or electronic stylus, or even via dictation). In everyday life, it is often these higher level affordances which guide decision making, because they are more directly tied to the goals that people are attempting to achieve. It might be useful to acknowledge this as outside the scope of the present investigation, yet still worthy of future investigation.

While this study focuses on the visuomotor nature of action affordances, where object suggest particular motor actions (e.g., pinch, clench, or poke), we acknowledge that actions based on affordance can be constructed at multiple levels of abstraction as pointed out by the reviewer. Based on that we added a paragraph into the discussion section as follows:

"It is noteworthy that action affordances operate across multiple levels of abstraction, extending beyond the visuomotor focus of our current study (Yi and O'Doherty, 2023). Consider how we interact with a keyboard: while it affords simple physical actions like key pressing, it also enables complex goal-directed behaviors like composing emails, coding software, or creating documents. Such higher-order affordances often guide our decision-making in daily life, as they connect directly to intended outcomes rather than to specific motor patterns. Although the present study focused on basic visuomotor affordances, the relationship between affordance and value-based decision-making investigated here may shed light on how abstract affordances influence choice behavior more broadly. Promising directions for future research include examining how abstract affordances might engage additional neural pathways beyond those revealed here."

2. Even basic motor affordances must often be learned. Again returning to the keyboard example, the knowledge that a keyboard affords "pressing" is tied to experience with that object – for example, knowing that the keys are spring-loaded. The same is true for many of the objects investigated in the present study. In future work, it might be helpful to use cross-cultural differences in affordance knowledge to distinguish between more fine-grained computational models. For example, an experienced user of chopsticks likely perceives far more affordances in them than someone who has never used them – and people of both types are likely available for recruitment in the authors' environs. This will allow future studies to disambiguate whether the affordances under study are habit like or driven by more innate processes.

We appreciate this comment about motor affordances being learned behaviors and the extent to which cross-cultural differences might influence affordance knowledge. While we attempted to control for this by selecting objects that predominantly afforded one specific action (pinch, clench, or poke), we recognize that participants' familiarity with objects could affect the underlying neural mechanisms, even if the final action affordances were similar. For instance, our data showed considerable variation in participants' familiarity with 'unfamiliar' objects like car parts or electronic music instruments, reflecting individual life experiences. Those familiar with such objects might process not just basic motor affordances but also complex functional affordances (e.g., car repair or music-making capabilities).



Although the size of our affordance dataset precluded analysis of familiarity-based differences, your suggestion to explore cross-cultural variations in affordance perception presents an excellent avenue for future research. Studying participants with diverse object-use experience could help determine whether affordances are primarily driven by habit, innate processes, or varying by the level of expertise. We have incorporated this perspective into our discussion section as follows:.

“While our study primarily examined basic motor affordances that were consistent across participants, we recognize that object familiarity may shape affordance perception in important ways. For individuals with specific expertise, objects such as car parts or electronic instruments might suggest sophisticated affordances (e.g., repair capabilities, music production) beyond the simple motor actions (pinch, clench, poke) we studied. Although we controlled for variability by selecting objects with strong statistical preferences for specific actions, the underlying neural mechanisms may differ based on familiarity. The current study was not designed to enable us to study how varying familiarity levels influence affordance perception. However, future research could address this by examining cross-cultural differences in affordance and/or by recruiting participants with diverse object-use experience. Such studies could help determine whether affordances stem primarily from learned habits or innate processes.”

3. Many objects offer not just action affordances, but also affective affordances. For

example, a can of beer offers not just the motor affordances of grasping or drinking, but also the affective affordance of drunkenness. The effects of such affective affordances – although rarely termed as such – has been extensively investigated in the substance use, addiction, and nutrition literatures within psychology. It would be helpful to make a connection between the present findings and those literatures, because I believe that they could benefit significantly from the conceptual clarity and computational specificity of the current approach (and this paper may receive more – deserving – attention as a result).

Thank you very much for your helpful and constructive suggestion. Based on your feedback, we discussed how our computational framework for action-affordance-based and value-based decision-making could be extended to affective affordances, particularly in the context of addiction, substance use, and related treatments. As a result, we have added the following paragraph to the discussion section.

"The findings of this study on action-affordance-based and value-based decision-making can be extended to the domain of affective affordances, offering valuable insights into substance use and addiction treatment. Relevant questions include how action affordances guide behavior by potentiating specific responses, affective affordances, such as cravings triggered by substances, elicit automatic emotional responses through learned associations. The meta-control mechanism described in the study, involving brain regions like the ACC and pre-SMA, aligns with cognitive control processes essential for regulating cravings and resisting substance-related cues in addiction (Everitt & Robbins, 2016; Goldstein & Volkow, 2011). This suggests that therapies targeting cognitive control, such as cognitive-behavioral therapy (CBT) or mindfulness-based interventions, could strengthen meta-control over affective affordances. Incorporating this dual-system arbitration framework into addiction treatment models may support adaptive and personalized intervention strategies (Kober et al., 2010)."

Reviewer #2 (Remarks to the Author):

The study by Yi and O'Doherty investigates the computational and neural mechanisms underlying the influence of action affordance on value learning. By combining functional neuroimaging, psychophysics, and computational modeling, they aim to understand how action affordance shapes value-based learning in novel environments. Their main finding suggests that action affordance operates as an independent system that concurrently guides action selection alongside value-based decision-making. These two systems engage in a competitive process to determine the best action. Notably, the authors demonstrate that different brain regions contribute to this process: the pre-SMA encodes the performance difference between the action-affordance and value-based decision-making systems, while the posterior parietal cortex (PPC) integrates predictions from both systems to facilitate action selection. Overall, this study provides important and novel insights into how the brain integrates information from multiple sources to optimize action selection. It fits within the scope of Nature Communications, but there are some points that the authors can address to further improve their study.

1. The authors describe how the brain integrates information about the affordance and the value of action. There are a number of previous studies that explored how the brain integrates information from multiple sources to make decisions. For instance, there are studies by Paul Cisek, Richard Andersen, Paul Glimcher and others (see below), which showed that the brain integrates dynamically information from multiple sources (such as value and effort/cost) when making decisions. Is it likely that the findings of the Yi and O'Doherty study reflects the integration of expected reward and effort cost to perform an action?

This will be also consistent with the concept of desirability value, according to which the posterior parietal cortex integrates information from disparate sources, such as expected reward and effort cost. I suggest that the authors describe the previous studies by Cisek, Andersen, Glimcher, Schrater and others and explain how different are their findings from previous studies that have already reported that the brain integrates information about the affordance of action (i.e., cost/effort) and the value associated with that action.

• Canaveral CA, Lata W, Green AM, Cisek P. Biomechanical costs influence decisions made during ongoing actions. J Neurophysiol. 2024 Jul 11. doi: 10.1152/jn.00090.2024. Epub ahead of print. PMID: 38988286.
• Cos I, Duque J, Cisek P. Rapid prediction of biomechanical costs during action decisions.

J Neurophysiol. 2014 Sep 15;112(6):1256-66. doi: 10.1152/jn.00147.2014. Epub 2014 Jun 3. PMID: 24899673.
• Christopoulos V, Bonaiuto J, Andersen RA. A biologically plausible computational theory for value integration and action selection in decisions with competing alternatives. PLoS Comput Biol. 2015 Mar 24;11(3):e1004104. doi: 10.1371/journal.pcbi.1004104. PMID: 25803729; PMCID: PMC4372613.
• Christopoulos V, Schrater PR. Dynamic Integration of Value Information into a Common Probability Currency as a Theory for Flexible Decision Making. PLoS Comput Biol. 2015 Sep 22;11(9):e1004402. doi: 10.1371/journal.pcbi.1004402. PMID: 26394299; PMCID: PMC4578920.
• Kable J, Glimcher P (2009) The neurobiology of decision: consensus and controversy. Neuron 63: 733–745. pmid:19778504

We appreciate the comment highlighting the need to contextualize our findings within the broader literature on neural decision-making integration. Key studies by Cisek, Andersen, Glimcher, Schrater, and others have established how the brain integrates multiple decision factors with the posterior parietal cortex (PPC) emerging as a critical hub for processing variables like reward, effort, and biomechanical costs.

Cisek and colleagues, for example, demonstrated rapid computation of biomechanical costs during ongoing movements (Canaveral et al., 2024; Cos et al., 2014). Andersen and Christopoulos developed computational frameworks showing how the PPC concurrently evaluates competing actions' value and effort, enabling real-time decision adjustment (Christopoulos et al., 2015; Christopoulos & Schrater, 2015). Glimcher and Kable revealed how primate brains encode value and choice across different regions, conceptualizing decision-making as a multistage process balancing multiple factors (Kable & Glimcher, 2009).

Our study extends this foundation by examining a novel aspect: the integration of action affordances with value-based choices during learning. We propose a distinct mechanism where affordance and value-based systems interact through a performance-based arbitration model when learning is involved. Unlike traditional approaches that treat affordance as a desirability factor (e.g., cost or effort), our model frames it as an independent decision-making system balanced against the value system through dynamically computed performance of value- and affordance-based systems. This performance functions as a context-sensitive meta-level controlling signal, with the prefrontal cortex (PFC) acting as the meta-controller to adjust each system's influence dynamically, while the PPC integrates this information for action selection.

While we initially considered modeling affordance as a cost factor, where incompatibility between value-based decisions and affordances would reduce action selection probability, this approach proved inadequate. When simulated using fitted parameters, the cost model predicted an increasing choice accuracy gap—contrary to observed data patterns (See Fig. 2 and Supplementary Fig. 8). This divergence from empirical results led us to develop our current performance-based arbitration model.

To further clarify these distinctions, we expanded our discussion with the following paragraphs to include a comparison with these studies:

"Our findings contribute to the understanding of decision-making by highlighting the dynamic interplay between affordance-based and value-based systems during learning-driven action selection. Previous research has established the PPC as a critical hub for integrating decision-related variables such as reward, effort, and biomechanical costs (Canaveral et al., 2024; Cos et al., 2014; Christopoulos et al., 2015; Christopoulos & Schrater, 2015). However, the present study furthers this understanding by demonstrating that the PPC's role goes beyond simply aggregating desirability factors. Instead, we propose a dynamic performance-based arbitration mechanism, where the PFC functions as a meta-controller, continuously adjusting the influence of affordance and value-based systems based on their predictive accuracy. The PPC then integrates these dynamically weighted inputs to guide action selection, suggesting a more complex and adaptive decision-making process.

This performance-based arbitration model helps address limitations observed in traditional cost-based frameworks in explaining our data. Previous models that conceptualized affordances as cost-like constraints, similar to biomechanical effort, appeared less effective when modeling learning-driven decisions. Specifically, simulations based on such models predicted an increasing accuracy gap inconsistent with our empirical data. To reconcile these discrepancies, we developed a model where affordance- and value-based systems operate as parallel processes moderated by performance-based meta-control from the PFC. This framework not only aligns with but also expands Glimcher and Kable's multi-stage decision-making theory (Kable & Glimcher, 2009). emphasizing that the PPC is not merely a passive integrator but an active decision-making arena modulated by adaptive, learning-sensitive signals from the PFC. This context-sensitive mechanism highlights how the brain dynamically balances multiple decision systems to optimize behavior."

2. It will be also interested in assessing if there is any temporal difference between the two

systems. Does the brain first assess the affordance of the action before the value or the other way? fMRI has a low temporal resolution, so it might not be possible to look temporal differences between the two systems, but it will be interested in trying to assess it.

Indeed fMRI does not allow us to address this question owing to its poor temporal resolution. However, to answer this question at a behavioral level it could be possible to experimentally manipulate the speed in which a decision needs to be made to see if speeded and non-speeded decisions are differently influenced by affordance and value. For instance, if value influences comes online later than the affordance-based process, then performance in incongruent conditions is likely to be poorer during quick decisions compared to slow decisions, even post-learning. This would suggest that value information takes time to override the dominant affordance-based initial response. Conversely, if affordances comes on line later than value-based mechanisms, the reverse outcome would be anticipated – participants should perform better (choosing the more valued action) during fast decisions and worse during slow decisions. The presence of a central arbitration module would imply no performance difference between fast and slow decisions throughout the learning period. However, since our study did not specifically employ a speeded decision-making paradigm, we instead conducted a supplementary analysis comparing trials with quick versus slow reaction times as an alternative approach.

For this we conducted GLM analyses on the incongruent trials (log_rt ~ (choosing_aff + choosing_correct + trial_idx + high_low + C(action)).

<Behavioral>

Mixed Linear Model Regression Results

Model:	MixedLM	Dependent Variable:	log_rt
No. Observations:	8682	Method:	REML
No. Groups:	19	Scale:	0.1333
Min. group size:	380	Log-Likelihood:	-3679.4712
Max. group size:	479	Converged:	Yes
Mean group size:	456.9		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	6.969	0.059	117.790	0.000	6.854	7.085
C(action)[T.pinch]	-0.022	0.019	-1.203	0.229	-0.059	0.014
C(action)[T.poke]	-0.016	0.013	-1.265	0.206	-0.041	0.009
choosing_aff	-0.034	0.014	-2.468	0.014	-0.061	-0.007
choosing_correct	-0.060	0.019	-3.101	0.002	-0.098	-0.022
trial_idx	-0.006	0.001	-4.992	0.000	-0.008	-0.003
high_low	-0.054	0.014	-3.927	0.000	-0.081	-0.027

<fMRI>

Mixed Linear Model Regression Results

Model:	MixedLM	Dependent Variable:	log_rt
No. Observations:	7171	Method:	REML
No. Groups:	30	Scale:	0.1462
Min. group size:	231	Log-Likelihood:	-3432.1411
Max. group size:	240	Converged:	Yes
Mean group size:	239.0		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	6.938	0.053	130.726	0.000	6.834	7.042
C(action)[T.pinch]	-0.032	0.014	-2.327	0.020	-0.058	-0.005
C(action)[T.poke]	-0.043	0.019	-2.203	0.028	-0.081	-0.005
choosing_aff	-0.026	0.016	-1.571	0.116	-0.057	0.006
choosing_correct	-0.056	0.017	-3.274	0.001	-0.090	-0.023
trial_idx	-0.003	0.001	-2.719	0.007	-0.006	-0.001
high_low	-0.077	0.012	-6.152	0.000	-0.101	-0.052

The coefficients from the GLM analyses showed that choosing a correct action and choosing affordance-compatible actions both were associated with shorter RTs, however, the coefficient of choosing_aff was only significant in the behavior only study not the fMRI study.

In summary, performance was better in the fast RT trials but the slow RT trials were not consistently biased to affordance across studies. Thus, we cannot conclude that affordance reliably acts as a bias for value-based decisions, nor that value acts as a bias to affordance-based decisions.

As the reviewer pointed out the fMRI data doesn't have the temporal resolution to study the temporal difference between the two systems. Instead, we conducted connectivity analyses which we detailed in addressing the following comment.

Related to this point, we included following paragraphs into line 378 of page 12 of the manuscript.

"Another question that arises from our findings is the temporal dynamics between affordance-based and value-based processes. We considered behavioral approaches to explore this question. A speeded decision-making paradigm, for instance, could reveal whether one system dominates earlier in the decision process. If value information emerges later than affordance-based processes, then performance in incongruent conditions should be poorer during speeded decisions compared to slower ones, even post-learning. Conversely, if affordance-based processes come online later, participants should perform better in speeded decisions and worse in slower ones in incongruent conditions. Alternatively, if a central arbitration mechanism governs the integration of these processes, no significant performance differences would be expected across speeded and non-speeded decisions in incongruent conditions.

Since our study did not employ a speeded decision-making paradigm, we instead conducted supplementary analyses to compare trials with quick versus slow reaction times. GLM analyses on incongruent trials revealed that both choosing correct actions and affordance-compatible actions were associated with shorter RTs. Notably, the coefficient for choosing affordance-compatible actions was significant only in the behavior-only study, but not in the fMRI study (Supplementary table 8; $\beta = -0.034$, $z = -2.47$, $p \leq 0.05$ for selecting affordance-compatible actions in the behavioral experiment, and $\beta = -0.026$, $z = -1.57$, $p = 0.116$ in the fMRI experiment). While performance was better in fast RT trials, slow RT trials were not consistently biased toward affordance-based decisions across studies. Therefore, we cannot definitively conclude that affordance acts as a consistent bias for value-based decisions or vice versa."

3. Can the author assess whether and how the brain encodes the competition between available actions? For instance, it would be interesting to explore differences in brain

activation between quick and slow responses. Previous studies have shown that reaction time reflects the level of competition between actions; that is, the longer it takes to respond, the stronger the competition between the available actions.

As noted in our response to comment 2, reaction time in this task reflects a combination of factors, including learned value (learned contingency), affordance-compatibility, task progress, and other influences. Consequently, it is challenging to attribute slower reaction times solely to higher levels of competition between affordance-compatible actions and high-value actions.

Our GLM analysis demonstrated that the coefficient for affordance-based choices was the only significant predictor of increased reaction times in one of the collected datasets. Therefore, instead of directly using reaction time as a measure of competition, we employed the absolute difference between model-generated parameters of value and affordance for chosen actions as a proxy for competition. This metric was subsequently included in our psychophysiological interaction (PPI) analysis.

However, when using a group-level value region as the seed, we did not observe any regions with selectively stronger connectivity under conditions of higher competition ($t(29) = -0.225$, $p = 0.823$ for the value region; $t(29) = 0.431$, $p = 0.669$ for the affordance region; $t(29) = -0.626$, $p = 0.536$ for the performance difference region; $t(29) = -0.578$, $p = 0.568$ for the value-based performance region; $t(29) = -0.255$, $p = 0.801$ for the affordance-based performance region).

While these findings do not provide conclusive evidence about how the brain encodes competition between available actions in this task, we acknowledge the importance of this question and plan to address it in future research.

Related to this analysis, we added the following sentences in the main manuscript line 485.

"Additionally, we attempted to identify changes in functional connectivity between the action selection region and other identified regions (chosen value, chosen affordance, performance of value, or performance of affordance) when there is a competition between actions, and presumably weighing more on the value-based choice for the final decision. However, we did not find a statistically significant change in functional connectivity. (For additional details, see Note 4 in the Supplementary Materials.)"

and a note in the supplementary materials as follows.

"**Note4: Examining functional connectivity changes during action competition**

We aimed to examine changes in functional connectivity when competing actions arise due to discrepancies between affordance-based and value-based choices. Rather than using RT as a proxy for competition, we employed the absolute difference between model-generated parameters of value and affordance for chosen actions. This decision was based on our GLM analysis, which demonstrated that RT is influenced not only by competition between actions but also by task progress and value. The absolute difference between these model-generated parameters was subsequently incorporated into our psychophysiological interaction (PPI) analysis.

However, when using a group-level value region as the seed, we did not observe any regions with selectively stronger connectivity under conditions of higher competition ($t(29) = -0.225$, $p = 0.823$ for the chosen value region; $t(29) = 0.431$, $p = 0.669$ for the chosen affordance region; $t(29) = -0.626$, $p = 0.536$ for the performance difference region; $t(29) = -0.578$, $p = 0.568$ for the value-based performance region; $t(29) = -0.255$, $p = 0.801$ for the affordance-based performance region)."

Minor points.

1. Please add line numbers in the manuscript. It is easier to review it.

Ok. We added the line numbers in the revised manuscript.

2. The authors need to improve the quality of writing in some parts of the manuscript. For instance, there are some parts that there is no space between the sentences (e.g., last line in page 6, "...experiment.Because...").

Ok. We revised any typos or mistakes in the revised manuscript.

Reviewer #3 (Remarks to the Author):

This study must have been a massively laborious undertaking, combining RL, mode-based fMRI, and automated decoding of gesture-based response. It is thus very unfortunate that I do not find the research question noteworthy. The interaction between action affordability and value-based decision-making is a very interesting topic, but the experimental paradigm in this study is incredibly artificial, at least for two reasons:

1. It is unclear if this task, in which participants see an object associated with (for instance) a pinching gesture and instead make a clenching gesture to earn a probabilistic reward, captures any real-world learning problems.

Thank you for your valuable feedback. We acknowledge your concern regarding the direct real-world applicability of our task. In real-world contexts, there are numerous examples where people must learn unintuitive gestures to use unfamiliar products or tools effectively.

One illustrative example is the use of chopsticks. While a stick naturally suggests a simple clench or pinch gesture depending on its size, mastering chopsticks involves learning a specific pinching gesture that is far removed from the basic action affordance for sticks. This adaptation is essential for achieving the rewarding outcome of successfully grasping food.

Additionally, this research sheds light on why certain products, like the iPhone or iPad, are remarkably easy to use. These devices are designed with compatible affordances that align with intuitive user interactions, thereby reducing the learning curve. Conversely, products that fail to tap into existing affordances pose greater challenges for users, necessitating more effort and adaptation.

We recognize that our task, similar to multi-armed bandit tasks, is a laboratory-based abstraction. Although these may not perfectly mirror real-world scenarios, they enable us to study human cognitive processes with exceptional precision. By abstracting these situations, we gain valuable insights into the mechanisms underlying behavioral adaptation in the presence of action affordances. These insights are crucial for understanding how people interact with technology and adapt to new tools, reflecting a real-world problem that our study addresses. This connection underscores the relevance of our findings to everyday challenges in learning and adaptation, influenced by the affordances present in the user environment.

We updated the last paragraph of the introduction as follows.

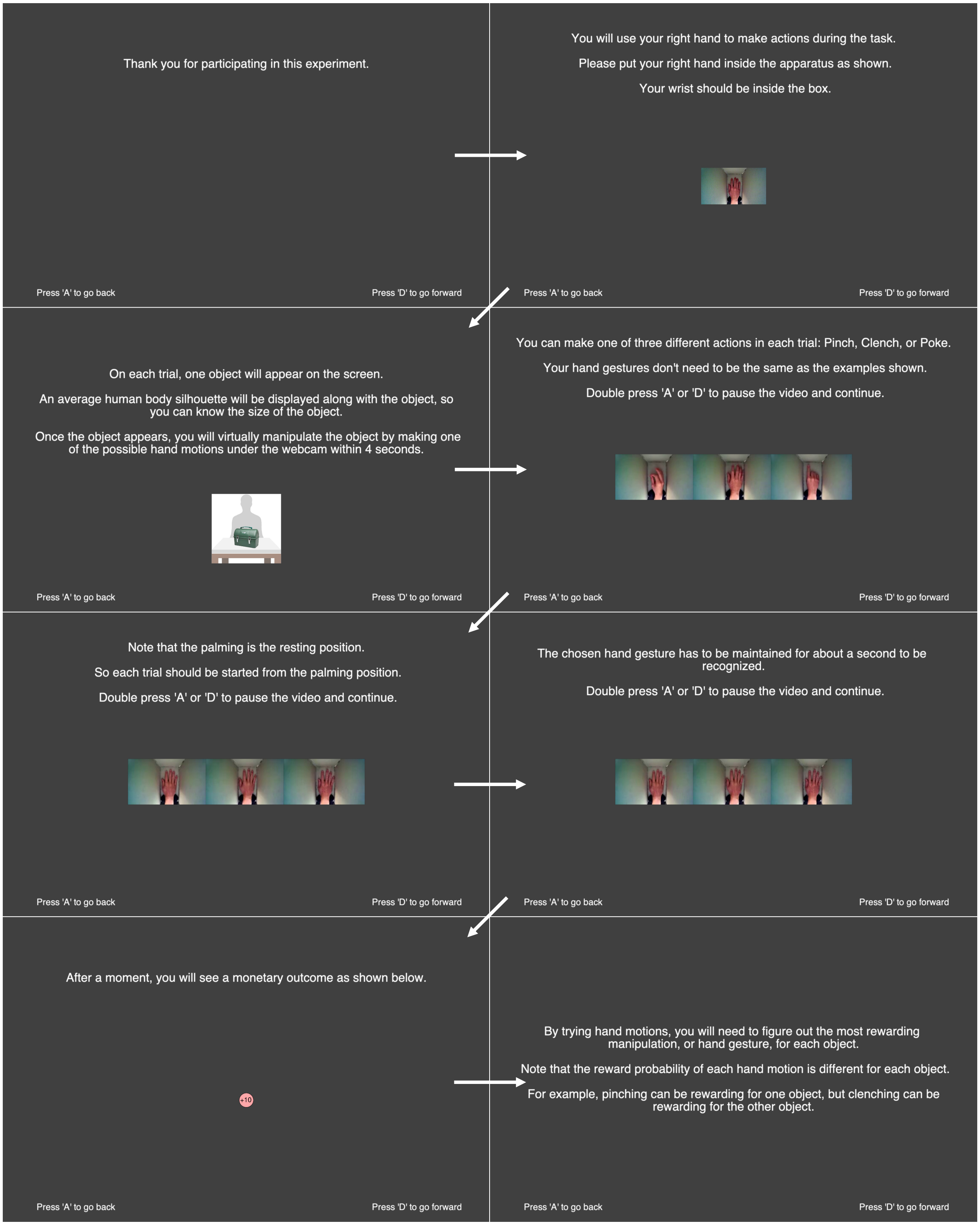
“To investigate how affordance influences decision-making, we designed a novel behavioral task that examines conditions in which the most rewarding action aligns with an affordance-compatible action and those in which they are incongruent. The incongruent conditions, in particular, capture a fundamental aspect of real-world action learning: the need to override intuitive action affordances when adapting to novel tools, interfaces, or environments. Many everyday tasks—such as learning to use chopsticks or mastering a musical instrument—require individuals to suppress immediate, affordance-driven responses in favor of newly learned actions that lead to the desired outcome.

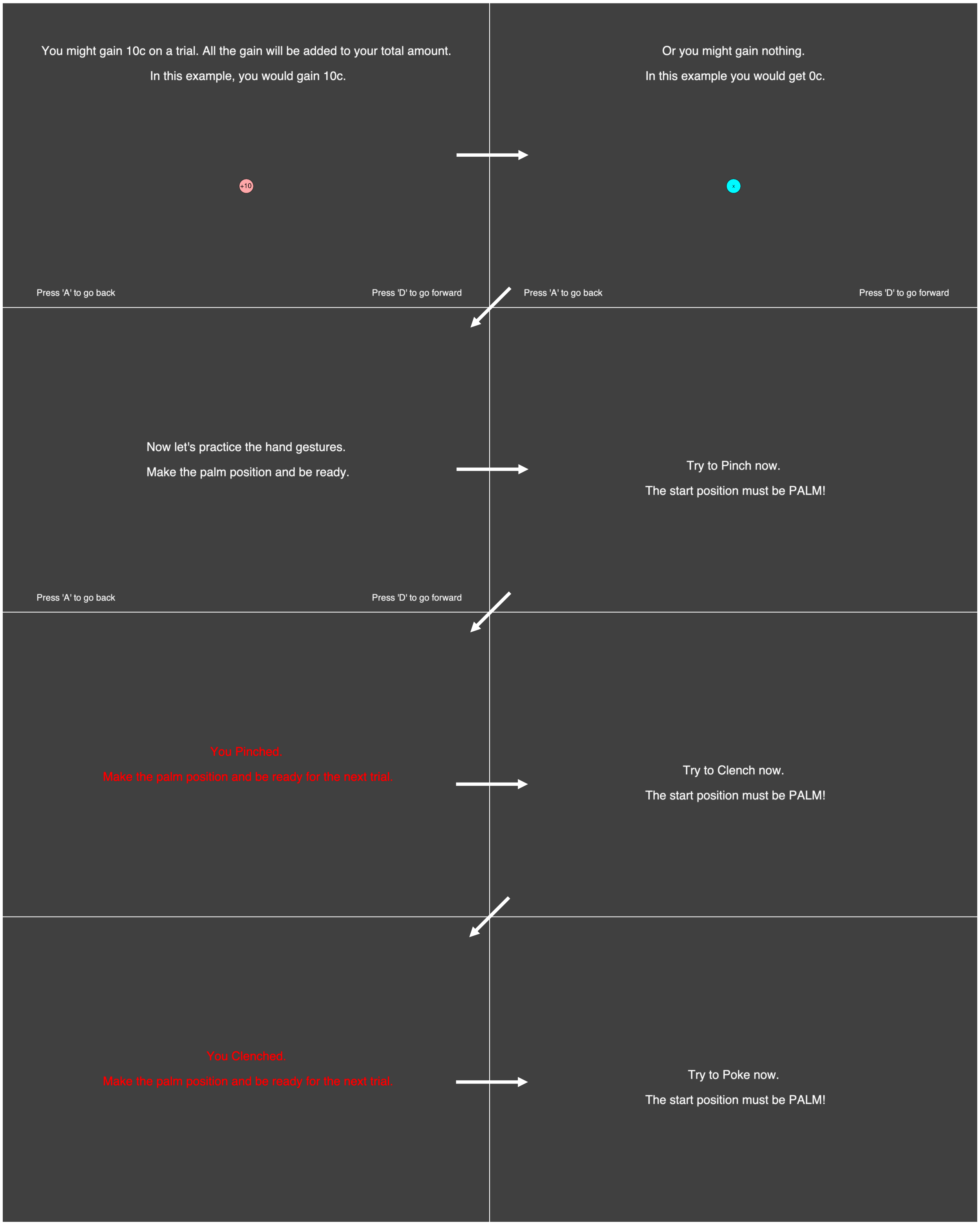
To systematically explore the mechanisms underlying this process, we employed a computational model-driven approach. We specified a series of computational models to capture different ways in which affordance might influence value learning and tested these models against human behavior. Additionally, we used functional MRI (fMRI) to measure brain activity and identify the neural correlates of these computational processes.

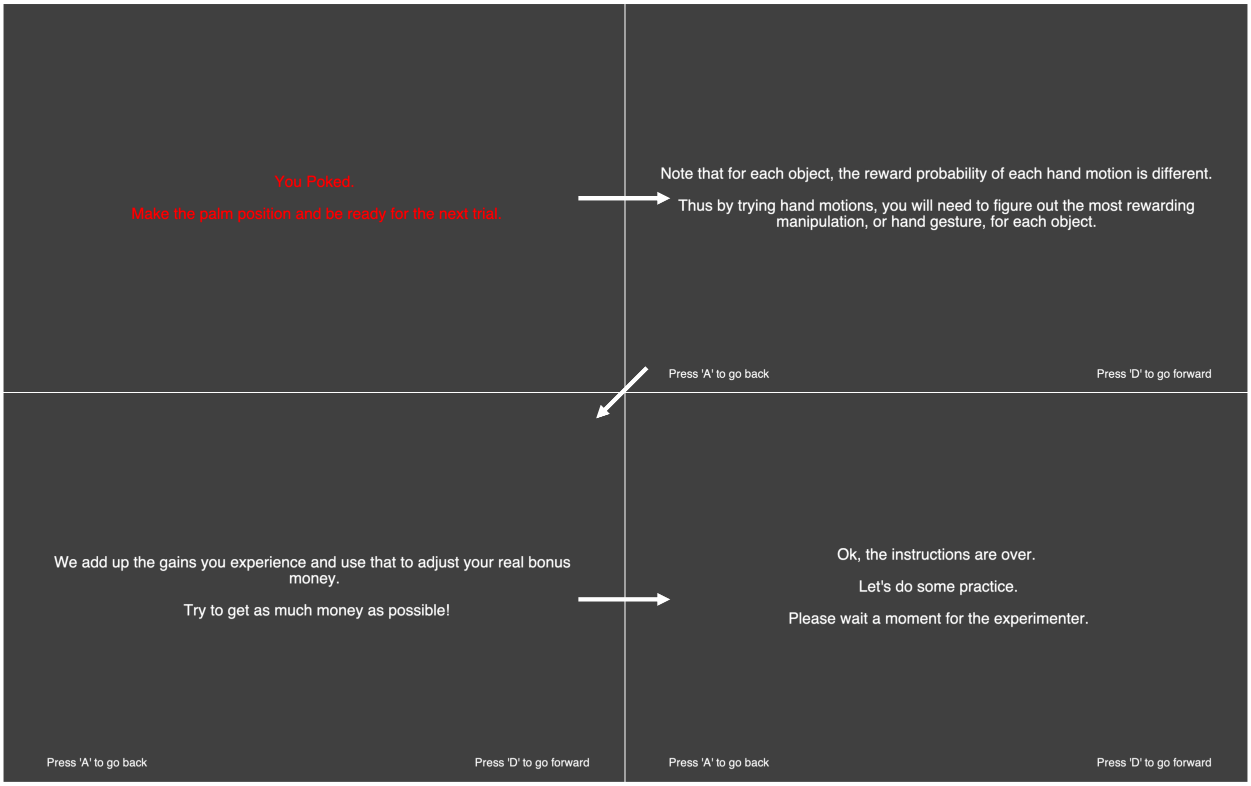
By examining how participants adjust their motor responses in response to probabilistic rewards, our study provides valuable insights into the cognitive mechanisms supporting this type of behavioral adaptation. Just as real-world learning involves trial and error, feedback processing, and reinforcement learning, our task allows us to isolate and analyze these processes in a controlled setting. Thus, the overarching goal of this study is to investigate how action affordances interact with value-based action learning across behavioral, computational, and neural levels.”

2. The experiment was divided into blocks, in each of which participants learned the most rewarding actions for four items, each corresponding to one of the four conditions (CH, CL, IH, and IL). The descriptions of instructions given to participants are entirely missing in this manuscript, but even if they were not told anything about reward contingencies, it is natural for them to eventually start trying to figure out which items are in the congruent or incongruent condition in each block (it is perhaps one of the most rational strategies they can take).

The instruction has been added to the supplementary figures 17, 18 and 19 as follows.


Thank you for participating in this experiment.
Press 'A' to go back
Press 'D' to go forward
You will use your right hand to make actions during the task.
Please put your right hand inside the apparatus as shown.
Your wrist should be inside the box.
Press 'A' to go back
Press 'D' to go forward
On each trial, one object will appear on the screen.
An average human body silhouette will be displayed along with the object, so you can know the size of the object.
Once the object appears, you will virtually manipulate the object by making one of the possible hand motions under the webcam within 4 seconds.
Press 'A' to go back
Press 'D' to go forward
You can make one of three different actions in each trial: Pinch, Clench, or Poke.
Your hand gestures don't need to be the same as the examples shown.
Double press 'A' or 'D' to pause the video and continue.
Press 'A' to go back
Press 'D' to go forward
Note that the palming is the resting position.
So each trial should be started from the palming position.
Double press 'A' or 'D' to pause the video and continue.
Press 'A' to go back
Press 'D' to go forward
The chosen hand gesture has to be maintained for about a second to be recognized.
Double press 'A' or 'D' to pause the video and continue.
Press 'A' to go back
Press 'D' to go forward
After a moment, you will see a monetary outcome as shown below.
+10
Press 'A' to go back
Press 'D' to go forward
By trying hand motions, you will need to figure out the most rewarding manipulation, or hand gesture, for each object.
Note that the reward probability of each hand motion is different for each object.
For example, pinching can be rewarding for one object, but clenching can be rewarding for the other object.
Press 'A' to go back
Press 'D' to go forward


You might gain 10c on a trial. All the gain will be added to your total amount.
In this example, you would gain 10c.
+10
Press 'A' to go back
Press 'D' to go forward
Or you might gain nothing.
In this example you would get 0c.
Press 'A' to go back
Press 'D' to go forward
Now let's practice the hand gestures.
Make the palm position and be ready.
Press 'A' to go back
Press 'D' to go forward
Try to Pinch now.
The start position must be PALM!
You Pinched.
Make the palm position and be ready for the next trial.
Try to Clench now.
The start position must be PALM!
You Clenched.
Make the palm position and be ready for the next trial.
Try to Poke now.
The start position must be PALM!



As shown in the new supplementary figures 17, 18, and 19, participants were not informed that half of the items have congruent actions as the most rewarding actions. Instead, they were simply instructed to identify the best hand gestures for each object, with no mention of the affordances.

It can also be observed that participants, unaware of the reward structure, consistently showed an initial exploration bias towards affordance-compatible actions across all blocks. This indicates that participants were unable to decipher the task structure in which half of the items were in congruent conditions and the other half were incongruent. If participants had learned the task structure, there would likely be a gradual increase in the initial exploration bias. However, as can be seen in supplementary figure 2 a and c, the initial exploration bias stayed around 0.5. Furthermore, the mixed linear regression of initial selection bias on block index showed non-significant slopes in both behavioral and fMRI datasets (Fixed-effect coefficients for the slope: $\beta = -0.005$, $z=-1.115$ and $p=0.265$ in the behavioral data, $\beta=-0.002$, $z=-0.232$ and $p=0.817$ in the fMRI data, please check the details in the following table).

<Behavioral data>

```
           Mixed Linear Model Regression Results
=============================================================
Model:              MixedLM  Dependent Variable:  p_affordance
No. Observations:   228      Method:              REML
No. Groups:         19       Scale:               0.0453
Min. group size:    12       Log-Likelihood:      2.5831
Max. group size:    12       Converged:           Yes
Mean group size:    12.0
-------------------------------------------------------------
                       Coef.  Std.Err.   z    P>|z| [0.025 0.975]
-------------------------------------------------------------
Intercept               0.461    0.049  9.431 0.000  0.365  0.557
block_index            -0.005    0.005 -1.115 0.265 -0.014  0.004
Group Var               0.032    0.000
Group x block_index Cov -0.001   0.000
block_index Var         0.000    0.000
=============================================================
```

<fMRI data>

```
           Mixed Linear Model Regression Results
=============================================================
Model:              MixedLM  Dependent Variable:  p_affordance
No. Observations:   180      Method:              REML
No. Groups:         30       Scale:               0.0555
Min. group size:    6        Log-Likelihood:      -13.0331
Max. group size:    6        Converged:           Yes
Mean group size:    6.0
-------------------------------------------------------------
                       Coef.  Std.Err.   z    P>|z| [0.025 0.975]
-------------------------------------------------------------
Intercept               0.459    0.037 12.380 0.000  0.386  0.531
block_index            -0.002    0.010 -0.232 0.817 -0.023  0.018
Group Var               0.012    0.000
Group x block_index Cov -0.000   0.000
block_index Var         0.000    0.000
=============================================================
```

We added the following sentence to the caption of Supplementary Fig.2
"The mixed linear regression of initial selection bias on block index showed non-significant slopes in both datasets (Fixed-effect coefficients for the slope: β= -0.005, z=-1.115 and p=0.265 in the behavioral data, β =-0.002, z=-0.232 and p=0.817 in the fMRI data, for the intercept: β = 0.461, z=9.431 and p<0.001 in the behavioral data, β =0.459, z=12.380 and p<0.001 in the fMRI data)"

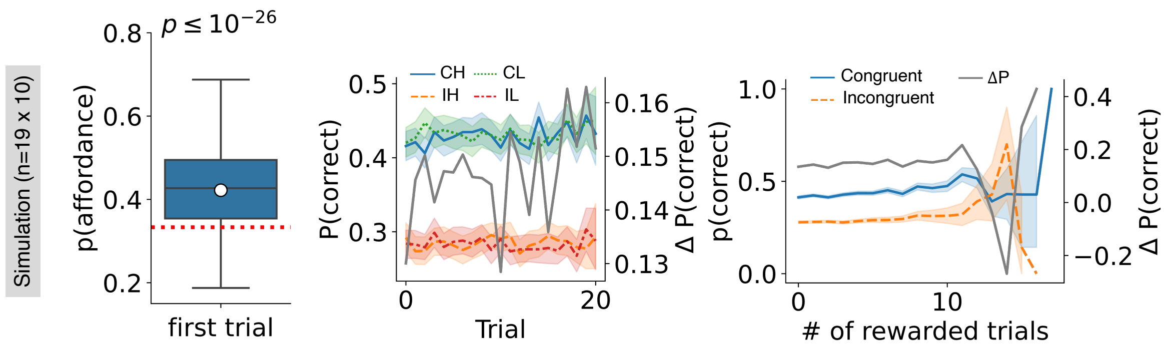
It is unclear the degree to which the reported findings are generalizable beyond this experimental paradigm. The point 2 is particularly crucial for the significance of the modeling results; the "performance-based arbitration model" is effectively equivalent to first testing if the most affordable action is the most rewarding one (which is the case 50% of the time) and trying the other two actions if not. The artificiality of the experimental paradigm heavily limits my enthusiasm for this manuscript. It does not help that the authors did not conduct much analysis to demonstrate that their complicated models characterize the behavioral data better than simpler, perhaps heuristic, accounts (such as the one I described above).

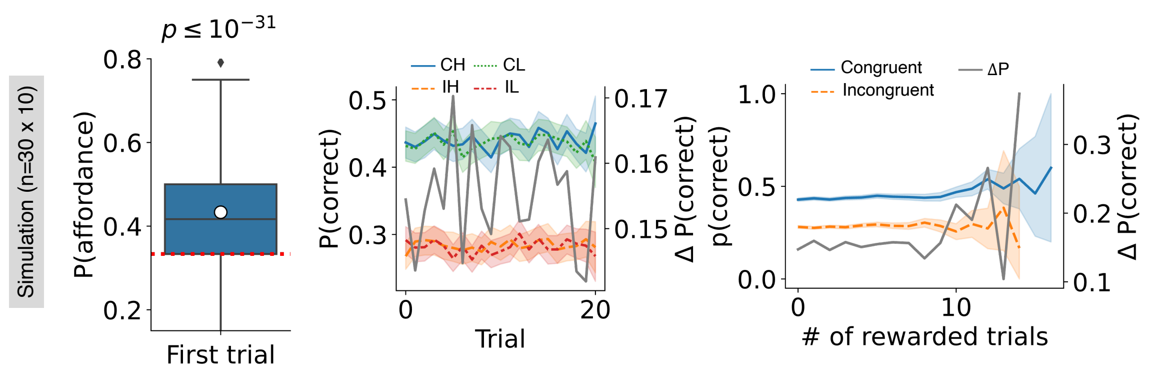
Thank you for your observations concerning the complexity of our proposed model. We appreciate the opportunity to clarify why a more complex model is necessary given the nuances of our experimental data.

Variability in Initial Choices: The experimental results demonstrate that participants did not consistently initiate their choices with affordance-compatible actions, contrary to the expectations of a simpler heuristic model. Specifically, Supplementary Figures 2b and d show that the frequency of initially selecting affordance-compatible actions was notably below 0.5 for half of the participants. This departure from a simple heuristic approach underscores the need for a model that can account for the observed variability in choices.

Implementation and Evaluation of Simpler Models: Responding to your suggestion for simplicity, we initially tested a win-stay-lose-shift strategy, presuming an initial preference for affordance-compatible actions. However, the model failed to align with the data, indicating that first choices were not always affordance-compatible. To address this, we adapted the model to include a softmax decision rule influenced by affordance compatibility (encoded with 1 or 0 for the simplicity) with an inverse temperature parameter (β), coupled with a win-stay-lose-shift strategy for subsequent choices.

However, the experimental data suggest that participants did not always prefer the affordance-compatible action as their initial choice, contrary to what might be expected under a simple heuristic model. Specifically, as demonstrated in Supplementary Figure 2b and d, the initial selection frequency of affordance-compatible actions was notably below 0.5 for half of the participants. This indicates variability in initial choices that cannot be adequately captured by a straightforward heuristic model.





Despite these modifications aimed at simplifying the decision rule, the simulations struggled to replicate the learning curves seen in the experiment, though they did accurately represent initial biases (see the above posterior predictive check figures). The log evidence of this simpler model was significantly poorer compared to more comprehensive models (behavioral data: log evidence = -18398; fMRI data: log evidence = -15317; see Figure 3 for log evidences of other models).

Necessity of a Complex Model: Consequently, our analyses suggest that the observed behavioral adaptations align more with a model incorporating gradual learning rather than simple heuristic strategies. Therefore, to computationally describe the reviewers' point on "first test affordance-compatible action and explore afterward", a gradual adaptation, or learning, is necessary with the affordance-based choice component. To capture this behavior, we employed models that incorporate elements of Bayesian or reinforcement learning alongside strategies that factor in affordance, such as bias, initial values, or distinct decision-making systems.

The "performance-based arbitration" model, in particular, effectively captures how strategies based on both affordance and value compete and adapt over time, in terms of model fitting metric and simulation. In response to the reviewer's concerns regarding the model's parameter specificity and the integration of affordance scores in decision-making, these aspects will be thoroughly addressed in the respective sections. As can be seen in the corresponding sections, this model has shown effective parameter recovery, and its adaptability to variations—whether through simplification of parameters or alternative methods of incorporating affordance—consistently reinforces the validity of our conclusions.

In conclusion, while we acknowledge the concerns regarding model complexity, our analysis supports the necessity of the arbitration mechanism in the model to adequately capture the intricate behaviors observed. Our findings suggest that simpler models, though appealing for their parsimony, fall short in explaining the depth and variability of human decision-making as demonstrated in our study.

Regarding this point, we updated "Note2: Additional model simulation results section" in the supplementary material as follows.

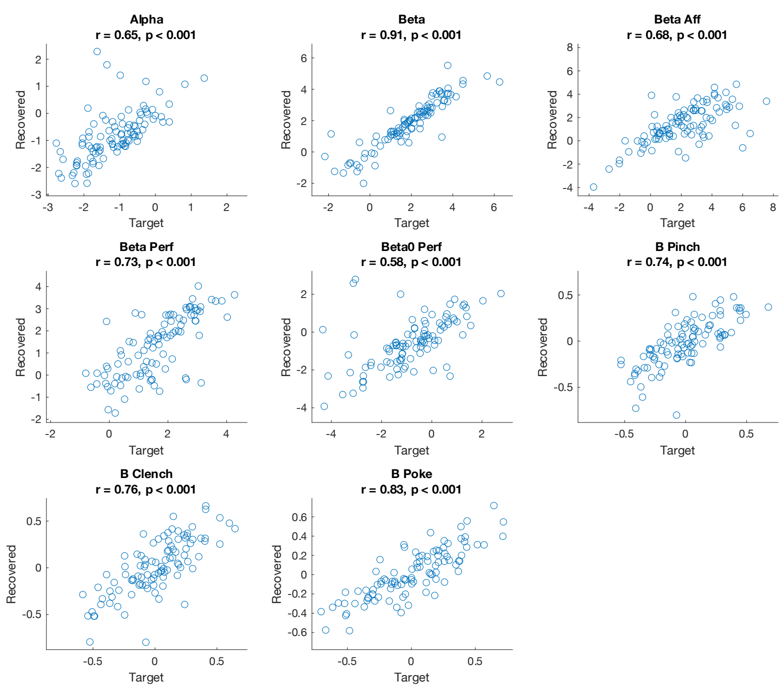
"Moreover, we implemented a WSLS-type model to further assess whether a simpler approach could account for the observed "first test affordance-compatible action and explore afterward" pattern in the data. This simplified model incorporated a softmax decision rule influenced by affordance compatibility (encoded as 1 or 0 for simplicity), modulated by an inverse temperature parameter, alongside a win-stay-lose-shift strategy for subsequent choices. While these modifications allowed the model to capture initial biases, it failed to replicate the learning curves observed in the experiment (see Supplementary Fig. 25). Furthermore, the log evidence for this simpler model was substantially lower compared to more comprehensive models, indicating a poorer fit to the data (behavioral data: log evidence = -18398; fMRI data: log evidence = -15317; see Fig. 3 for log evidence of other models)."

Below are some of the specific issues I had with the modeling approach:

3. The models seem overparameterized. Just an example -- even in the simplest RL model, the authors had b_pinch, b_clench, and b_poke as free parameters. They cannot be uniquely estimated. Forgive my frankness, but I am surprised that the authors did not catch this, and such a mistake casts doubts on the entire modeling results. One suggestion would be to make the code publicly available at this point for transparency.

Thank you for the valuable feedback and we acknowledge your concerns regarding potential overparameterization in the model, specifically regarding the parameters b_pinch, b_clench, and b_poke. Below, we address your points and provide evidence to support the robustness of our modeling approach.

Parameter Identifiability: To address this, as detailed in the Methods section, we conducted a parameter recovery analysis to assess identifiability given our task design of 480 trials per participant (960 for the behavioral experiment) by simulating the performance-based arbitration model 100 times with a wide range or free parameters. The analysis yielded an average Pearson correlation of 0.76 across the 8 parameters, indicating strong parameter recovery. The following plot demonstrates recovery results for each parameter, and no evidence of overparameterization was found under this experimental design.

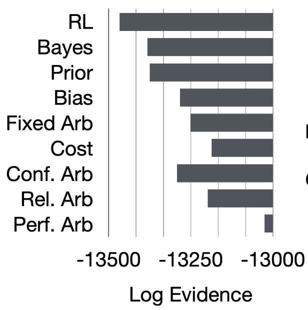


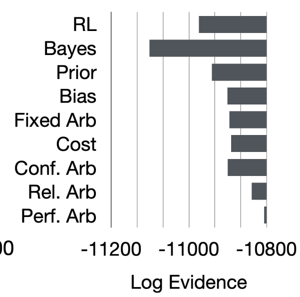
Alternative Parameterization: To further address the concern, we explored an alternative parameterization, where the action bias terms were reduced to two free parameters:

$b\_clench - b\_pinch = \beta_1$ and $b\_poke - b\_pinch = \beta_2$.

Despite this reduced parameterization, the performance-based arbitration model remained the best-fitting model.

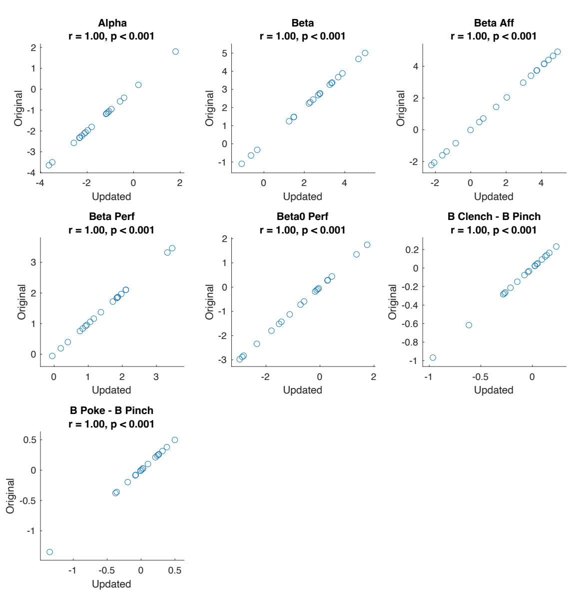
<Behavioral (n=19)> <fMRI (n=30)>

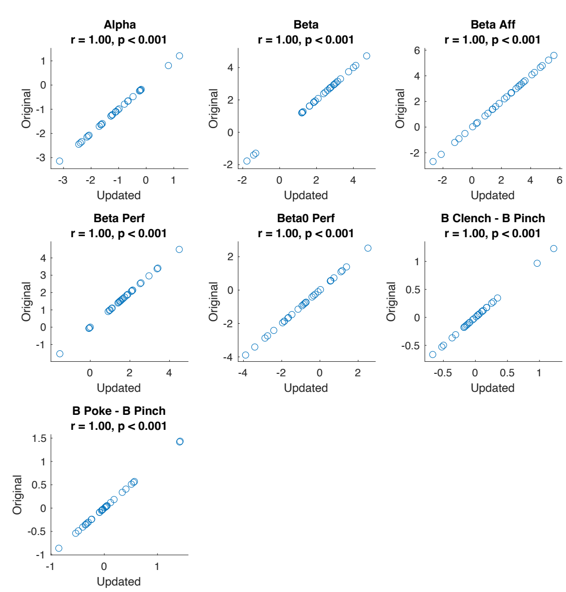




Additionally, we compared the fit parameters between the original and updated models, and the correlations (see the plot below) show that the two parameterizations produced highly consistent results.

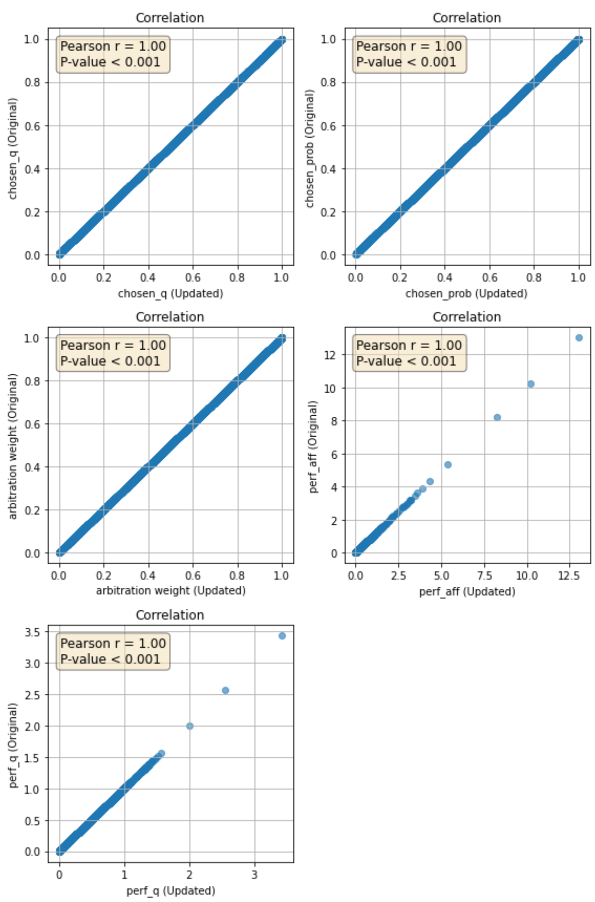
<Behavioral (n=19)> <fMRI (n=30)>



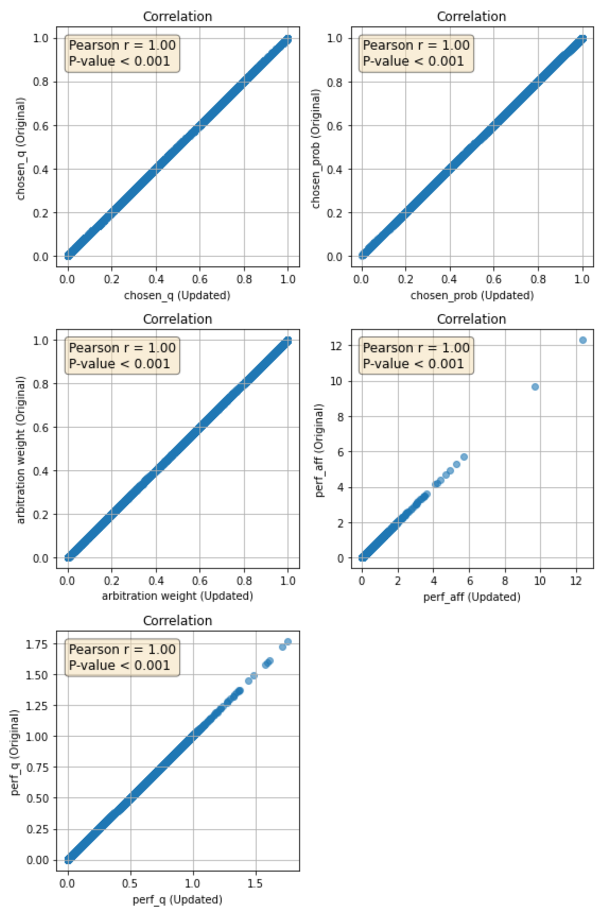


Collectively, the parameter recovery analysis and the results from the alternative parameterization demonstrate that while the original parameterization may theoretically raise concerns about identifiability, the task design and data volume ensured robust parameter estimation. Furthermore, the model-based cognitive variables derived from the reduced-parameter model were identical to those from the original model, reinforcing the robustness of our findings (see the plot below).

<Behavioral (n=19)>



<fMRI (n=30)>



To facilitate further evaluation and transparency, we have made our full codebase publicly available at https://github.com/sanghyunyi/Affordance2025. The repository includes all computational models and supporting analyses, which we hope will aid in understanding and verifying the modeling process and results.

Regarding this further analysis, we included a new Note 3 in the supplementary materials as follows.

“To evaluate concerns regarding potential overparameterization in our model—specifically related to the parameters b_pinch, b_clench, b_poke—we conducted a thorough parameter recovery analysis and explored an alternative parameterization. We acknowledge that, in theory, these parameters may not be uniquely identifiable in isolation. However, as detailed in the Methods section, we performed a parameter recovery analysis using a simulation-based approach. By generating synthetic data with the performance-based arbitration model across 100 simulations and a broad range of free parameters, we assessed the model’s ability to recover its true parameter values. Given our experimental design of 480 trials per participant (960 for the behavioral experiment), the analysis revealed an average Pearson correlation of 0.76 across the eight parameters, demonstrating strong parameter recovery. The Supplementary Fig. 26 illustrates the recovery results for each parameter, with no evidence of overparameterization under this task design.

To further validate the robustness of our modeling approach, we tested an alternative parameterization by reducing the action bias terms to two free parameters $b\_clench - b\_pinch = \beta_1$ and $b\_poke - b\_pinch = \beta_2$. Despite this simplification, the performance-based arbitration model continued to outperform other models, confirming its superior fit to the data in both behavioral and fMRI experiments (See Supplementary Fig. 27). Moreover, comparisons between the fit parameters and model-derived cognitive variables from the original and reduced models demonstrated perfect correlations (Pearson r=1.0), indicating that both parameterizations produced consistent results (See Supplementary Fig.28).”

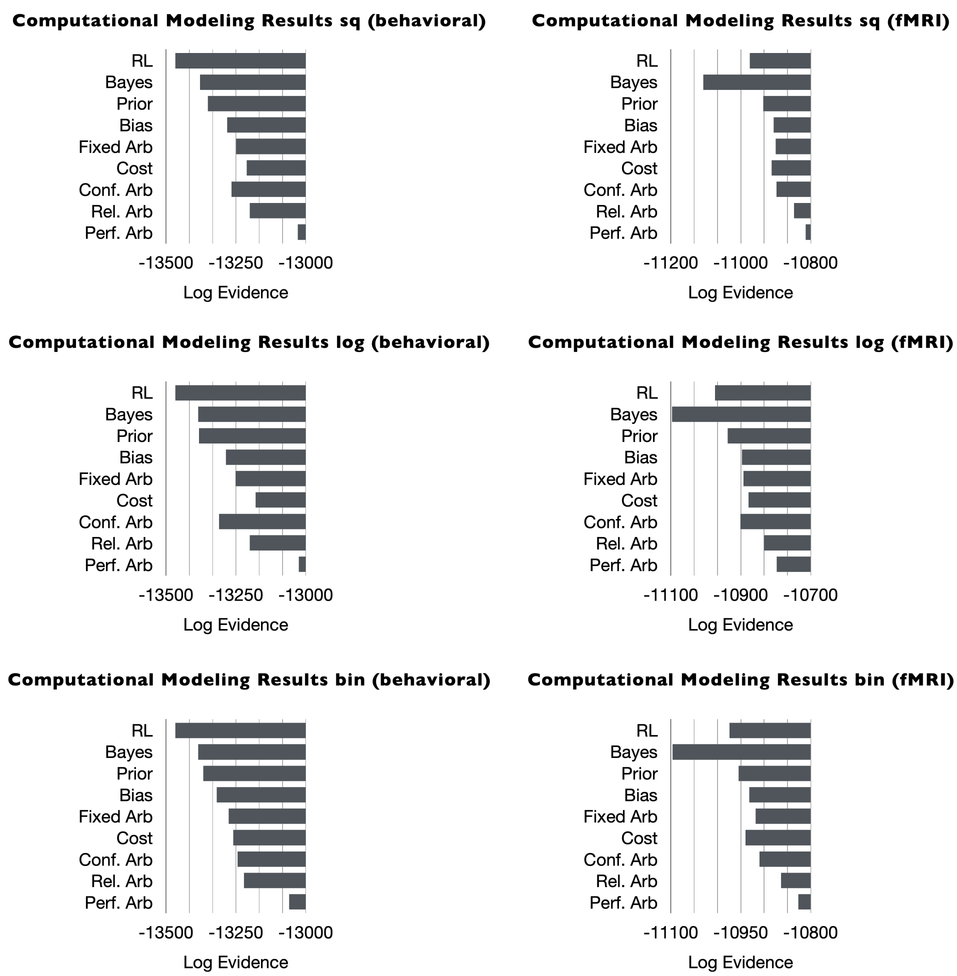
4. The authors make a lot of specific parametric assumptions in modeling, for example how affordance compatibility is linearly scored based on the post-task survey data. It is unclear to what degree the model comparison results depend on these (mis)specifications.

We appreciate the reviewer’s concern regarding the parametric assumptions made in modeling, specifically how affordance compatibility is linearly scored based on the post-task survey data. To address this, we tested three additional parameterization methods for affordance compatibility:

1. Squared affordance scores,
2. Log-transformed affordance scores, and
3. A binarized version, where only the maximum affordance score of pinch, clench, and poke is parameterized as 1, while the scores for other actions are set to 0.

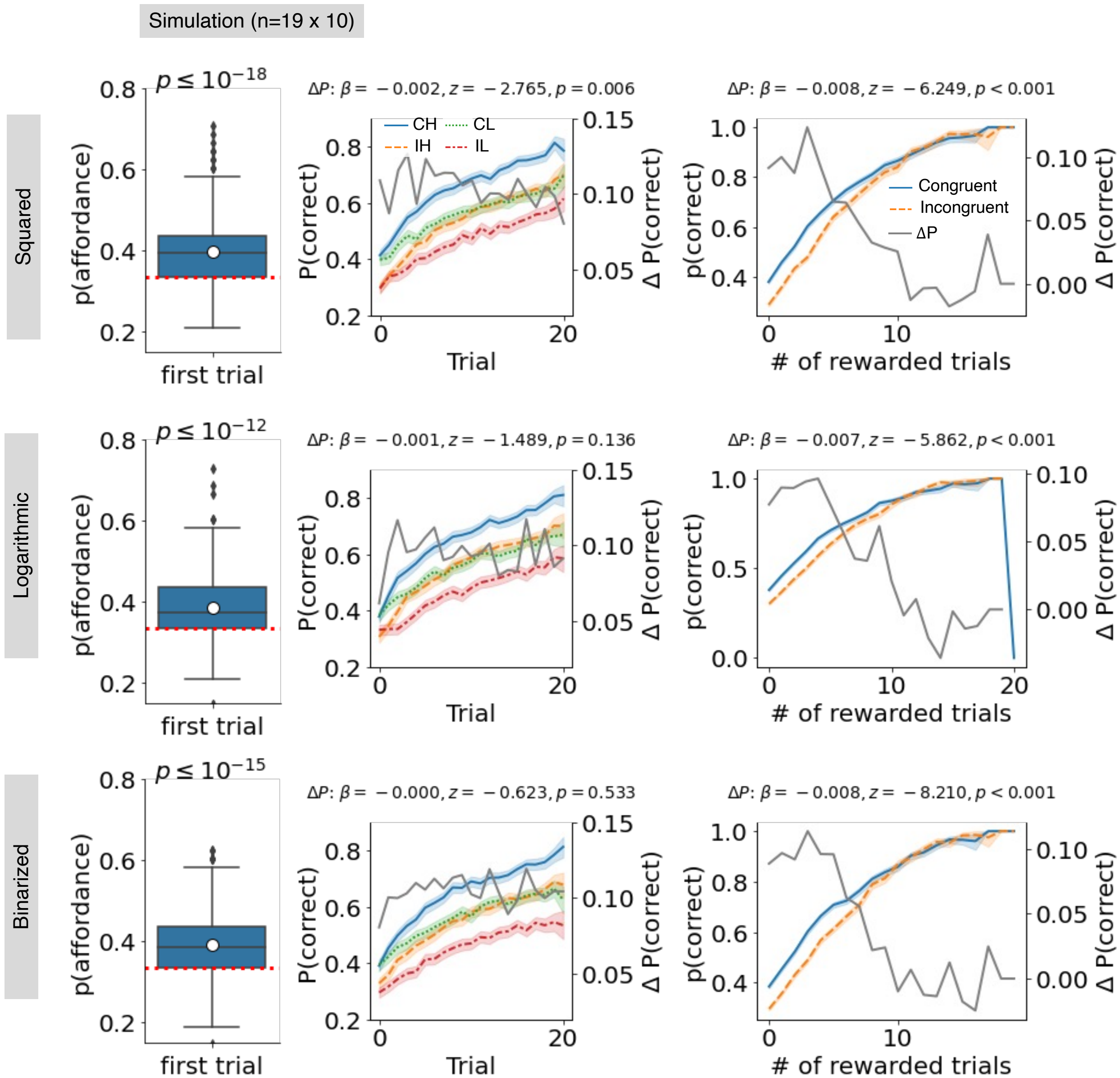
Additionally, in response to the reviewer's third point, we incorporated two free parameters to model action biases in these new sets of models.

Our results remain consistent across all three alternative parameterization settings. For instance, we still observed that the performance-based arbitration model achieved the best model fitting performance across both behavioral and fMRI datasets (see the following graphs).

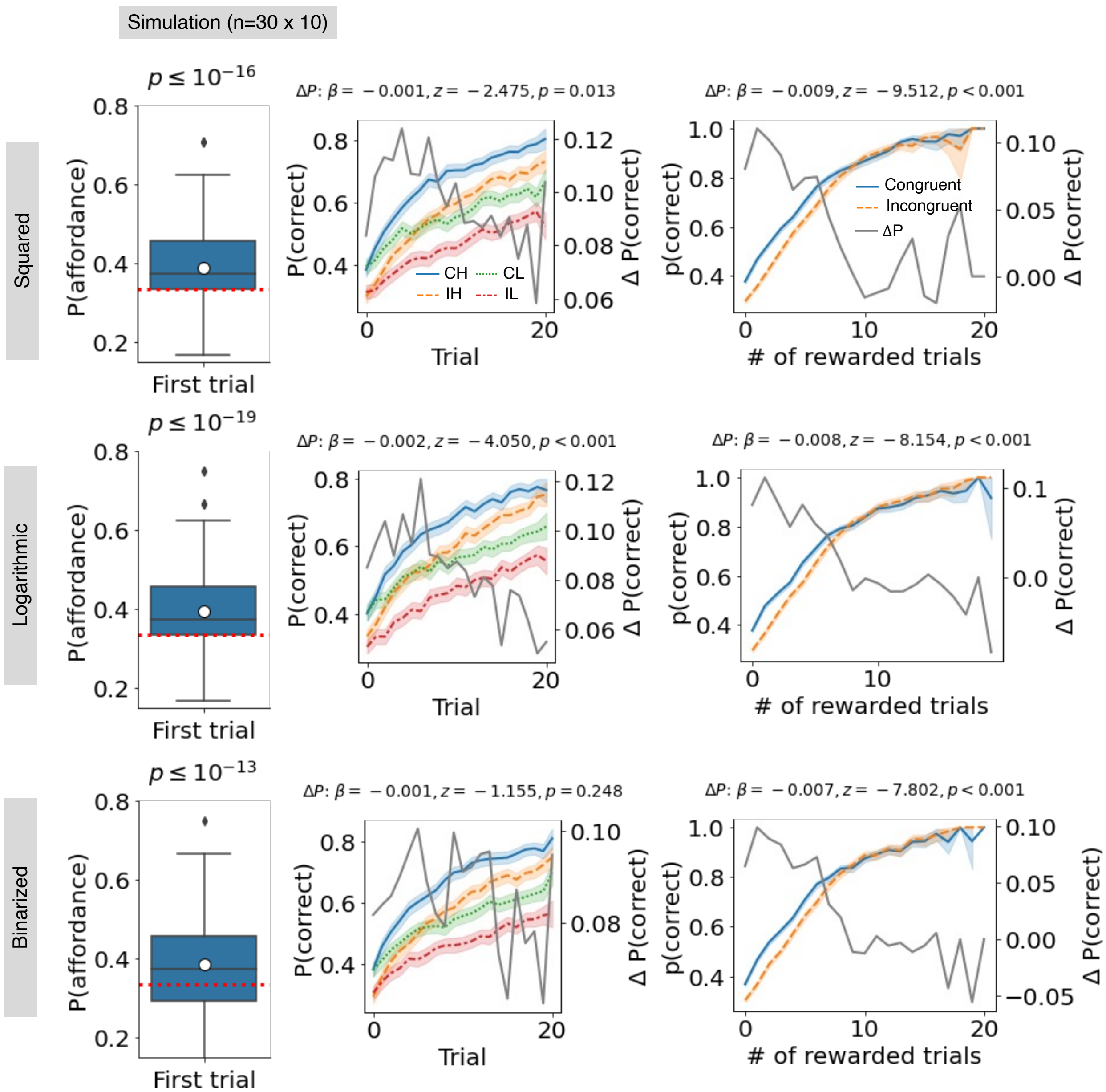


Furthermore, posterior predictive checks using the three types of performance-based arbitration models successfully captured the observed patterns of initial exploration bias, learning curves, and slopes. The results are consistent with the statistical tests reported in the manuscript, where we regressed the performance difference between congruent and incongruent conditions (ΔP) on the number of trials or the number of rewarded trials using the same mixed-effect GLM (see the attached plots with beta coefficients, z-statistics, and p-values).

<Behavioral simulation>



<fMRI simulation>



Based on these findings, we can confidently conclude that our primary conclusion that the performance-based arbitration model is the best is robust and not sensitive to the specific parameterization of the affordance score. We included the following subsection in the methods section.

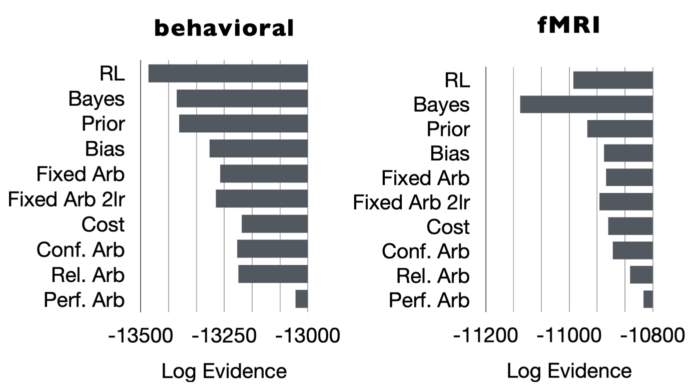
“**Alternative parameterization of affordance compatibility**

To address potential concerns regarding the parametric assumptions of affordance compatibility, we tested three additional parameterization methods based on post-task survey data: (1) squared affordance scores, (2) log-transformed affordance scores, and (3) a binarized version where the maximum affordance score of pinch, clench, or poke was set to 1, and all other scores were set to 0.

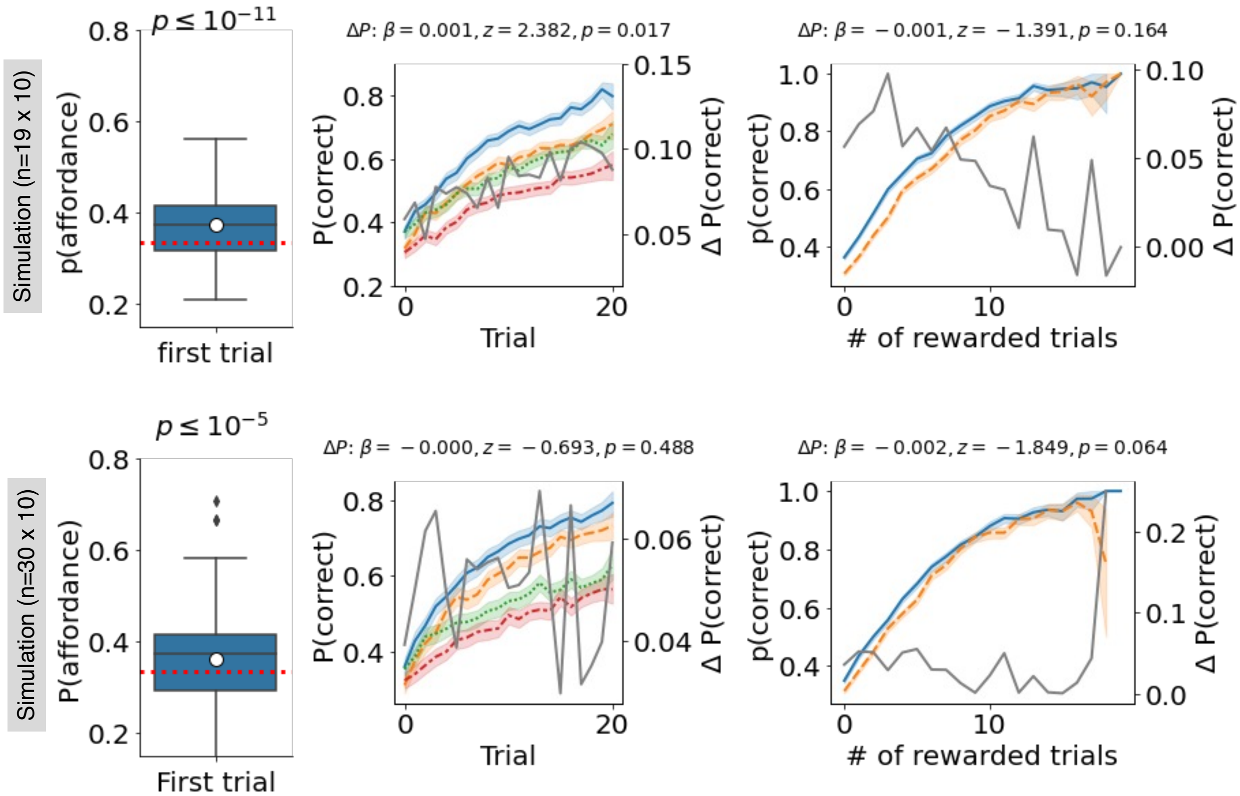
Consistent results across all parameterizations confirmed that the performance-based arbitration model provided the best fit to both behavioral and fMRI datasets (See Supplementary figure 20). Posterior predictive checks were conducted as well using the three performance-based arbitration models with alternative parameterizations. These models successfully captured observed patterns such as initial exploration bias, learning curves, and slopes. Statistical analyses, including mixed-effects GLM regressions of ΔP(performance difference) on trial progression and rewarded trials, further supported the robustness of the model's predictive accuracy and its insensitivity to specific parameterization choices (See Supplementary figures 21 and 22)."

5. The authors examined an alternative model, in which learning rates differ between congruent and incongruent conditions. They rejected it solely based on the lack of population-level difference in learning rate estimates, but it would make much more sense to reject it based on its performance (e.g. log evidence).

Thank you for pointing this out. Indeed, we conducted additional analyses to compare the models based on their performance. Specifically, the two learning rate models (2lr) exhibited worse log evidence compared to the single learning rate model, as shown in the provided log evidence graph, which includes the two learning rate fixed arbitration model.



Furthermore, the posterior predictive checks for the two learning rate models revealed their inability to capture the decreasing performance gap between congruent and incongruent conditions.

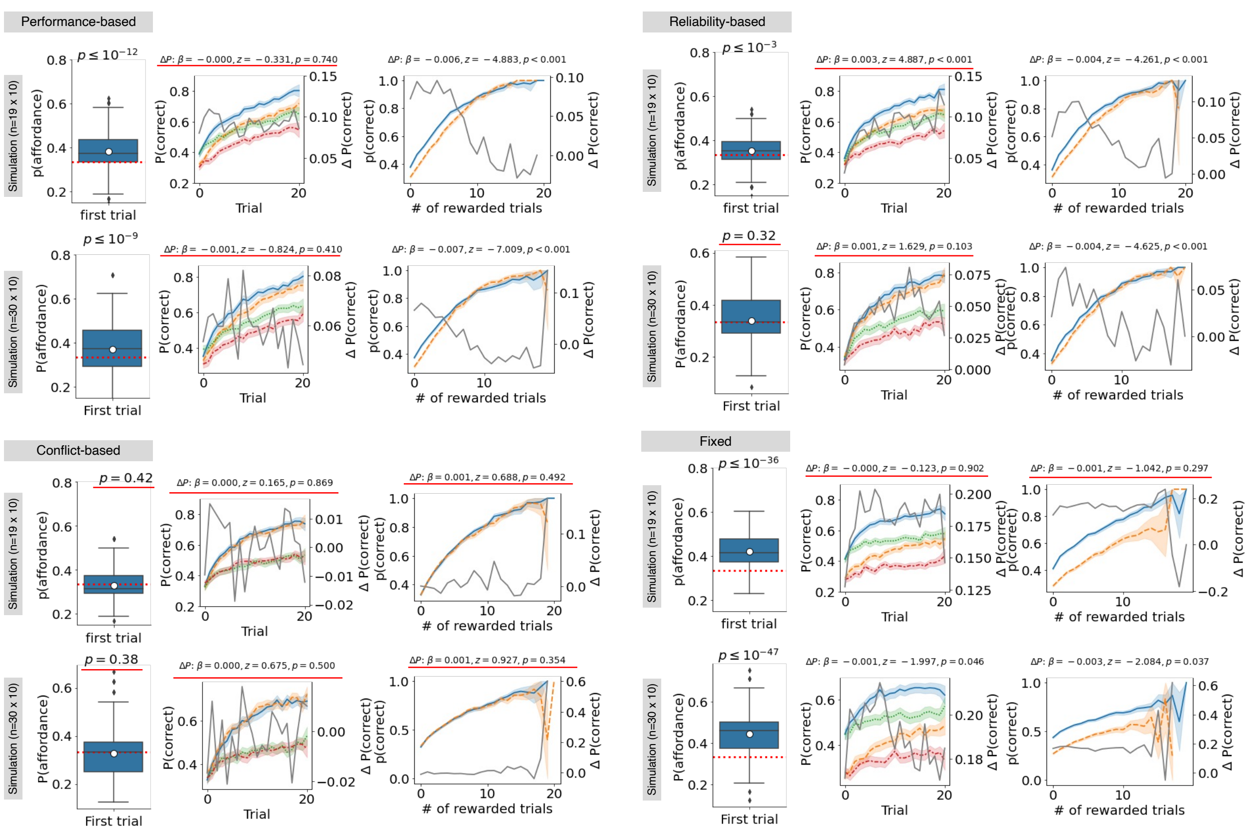


Based on both model evidence (log evidence) and posterior predictive checks, we found sufficient grounds to reject the two learning rate model and added the following sentence to the manuscript.

“Additionally, the model with two distinct learning rates exhibited lower log evidence compared to the single learning rate model and was unable to replicate the decreasing performance gap observed between congruent and incongruent conditions (see Supplementary Fig. 23).”

6. The authors wrote: "We didn’t use the hierarchical Bayesian inference function of the toolkit because model simulations using the hierarchically estimated parameters couldn’t reproduce the characteristics of the real data" (p.24). This sentence is vague (what characteristics?), but I am not convinced that it is a scientifically sound and unbiased reason not to use the hierarchical model, particularly given the common observation in the literature that it tends to help estimate parameters more reliably.

Thank you for your comment regarding the use of hierarchical Bayesian inference and for pointing out the need to provide further details. First of all we would like to highlight that in our experiment, each participant completed either 480 or 960 trials—approximately 5 or 10 times more trials per participant than the data sets commonly found in studies demonstrating the effectiveness of hierarchical Bayesian inference (Ahn et al., 2013; Piray et al., 2019). While hierarchical inference is highly valuable in situations where participant-level data are sparse, its utility diminishes when there are sufficient data points for reliable individual parameter estimation. In our case, we observed cases when the hierarchical approach introduced additional noise from other participants rather than improving parameter estimates. For example, in parameter recovery analysis using 100 times simulated data, the average correlation between recovered parameters of Performance-based arbitration model using the hierarchical Bayesian inference was 0.73 but without the hierarchical inference 0.76. This resulted in failures to reproduce the actual behavioral patterns in the posterior predictive checks (See the following figures. Statistical test results with a red underscore indicate failed cases).



For further context on the trade-offs of hierarchical Bayesian inference as a function of trial count, please refer to Figure 5 in Piray et al. (2019).

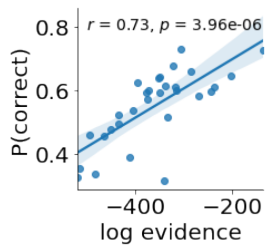
We hope this additional explanation clarifies our rationale for not using hierarchical Bayesian inference in this study and demonstrates that our decision was justified and tailored to the specific characteristics of our dataset. Accordingly, we updated the corresponding sentence into as follows.

“We did not utilize hierarchical parameter estimation in this study and relied instead on estimating individuals subject-level parameters. This was done because while hierarchical Bayesian inference for parameter estimation is especially useful in situations where participant-level data are sparse, its utility diminishes when there are sufficient data points for reliable individual parameter estimation (Ahn et al., 2013; Piray et al., 2019). In our case, each subject completed 480 or 960 trials, allowing for reliable estimation of individual subject-level parameters. In additional analyses we found that the hierarchical approach introduced additional noise from other participants into the individual subject-level estimates resulting in less accurate rather than more accurate parameter estimates. For example, in parameter recovery analysis using 100 times simulated data, the average correlation between recovered parameters of Performance-based arbitration model using the hierarchical Bayesian inference was 0.73 but without the hierarchical inference 0.76. This resulted in failures to reproduce the actual behavioral patterns in the posterior predictive checks (See Supplementary Fig. 24)”

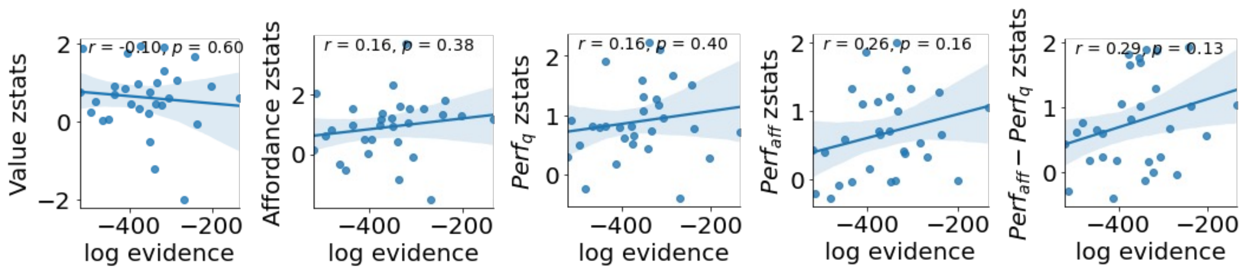
7. In fMRI analysis, the wrote wrote: "we found positive correlations between each participant’s propensity to choose the most rewarding actions and the extent to which BOLD activity in the identified arbitration regions can be explained by our model-derived arbitration variables" (p. 15). Doesn't it simply mean that, for those whose behavioral data was well characterized by a specific model, their MRI data was also well characterized by the same model?

Thank you for your observation regarding our interpretation of the fMRI findings. To clarify, our analysis revealed a correlation between participants' task performance and the extent to which their neural data could be explained by our model-derived arbitration variables. However, we did not find direct evidence to support the interpretation that participants whose behavioral data were well characterized by the performance-based model also exhibited neural data that closely aligned with the same model.

To further test this, we examined the relationship between the log evidence of the performance-based arbitration model and participants’ behaivoral task performance. For this analysis we indeed found a strong positive correlation, indicating that participants who were better explained by the performance-based arbitration model also performed better on the task.



However, when we turned to the fMRI data we did **not** observe a significant correlation between the log evidence of the performance-based arbitration model and the extent to which BOLD activity was explained by the same model-derived variables.



Thus, we hesitate to conclude that behavioral data well-characterized by the performance-based arbitration model directly implies that fMRI data will also be well-characterized by the same model. Instead, our findings suggest that arbitration between value-based and affordance-based action selection is crucial for achieving high task performance. Furthermore, our neural data support this idea, showing that individuals who utilized arbitration-related brain regions more effectively demonstrated better task performance. Thus, we'd like to reserve the interpretation that when the behavioral data well explained by the performance-based arbitration model, their fMRI data as well. Instead, we'd like to keep that arbitration between value-based and affordance-based action selection is necessary for achieving the high performance and the neural data further support the idea by showing that the people who are utilizing arbitration region for the arbitration showed better performance.

Reviewer #2 (Remarks to the Author):

The authors have thoroughly addressed all the key points raised in my review, and I appreciate their thoughtful revisions. I have no further concerns and recommend the manuscript for publication.

We sincerely thank the reviewer for their thoughtful feedback throughout the review process and for their positive evaluation of our revisions. We greatly appreciate the time and effort they invested in improving the quality of the manuscript.

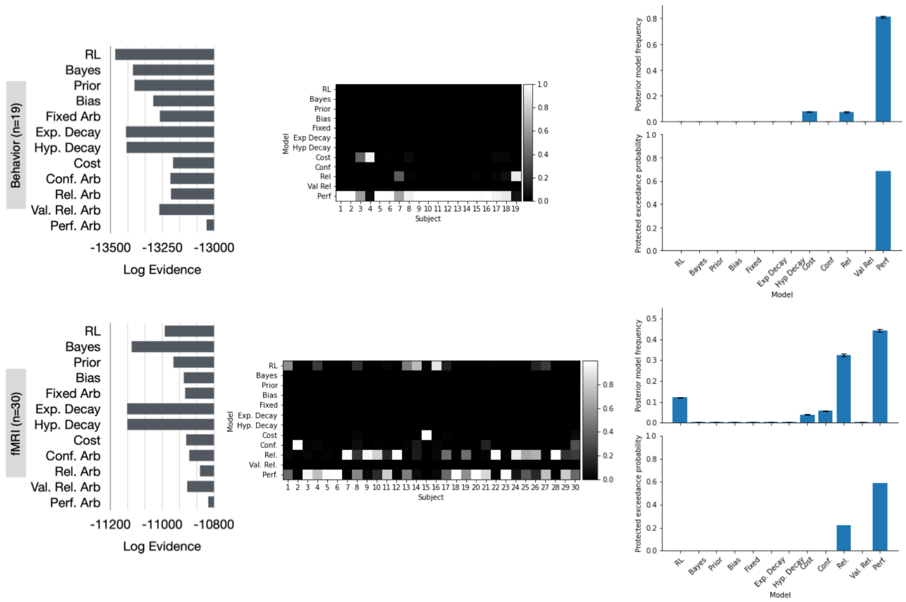
Reviewer #4 (Remarks to the Author):

In this paper the authors explore how reward information and action affordances co-shape value-based learning.
This is an interesting research question, which also has applied implications.
The reported experiments are well designed and reported; and the same holds for most analyses in the manuscript.

My main concern is whether the authors have sufficient evidence to back up their main computational claim. Indeed, the behavioural data rule-out computational models in which the action affordance bias is fixed over time. The authors proceed and propose a series of models in which a dynamic bias emerges from cognitive control mechanisms. However, can the authors rule out the possibility that the affordance bias has its own (passive/ decaying) temporal dynamics? Incorporating a time-varying bias could potentially "patch-up" the simpler models the authors tried out. Where is the evidence that there is an active control mechanism causing this time-varying bias? Perhaps there are certain qualitative aspects of the behavioural data that would lend support to the arbitration models over a time-varying bias. In some, there seems to be another class of models in-between the fixed and arbitration ones that the authors need to assess.

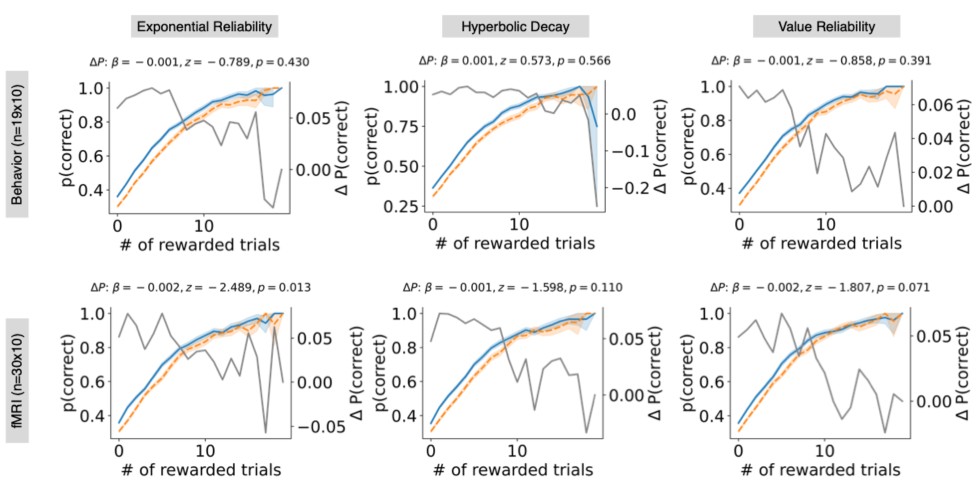
We thank the reviewer for highlighting this important theoretical alternative. To directly address this point, we implemented and tested additional models incorporating a passively decaying affordance bias. Two forms of decay, exponential and hyperbolic over trials, were examined.

However, when compared to our arbitration models, these passive decay models performed significantly worse in terms of model evidence and Bayesian model comparison. As shown in Supplementary Figure 29, models with passive decay consistently underperformed relative to those implementing arbitration via active control mechanisms.



<Supplementary Figure 29>

Furthermore, the passive decay models failed to capture key behavioral signatures that were successfully reproduced by the best-performing arbitration model, most notably, the decreasing gap between congruent and incongruent conditions in the frequency of correct action selection as a function of rewarded trials. Simulated results from both decay variants did not show a significant reduction in the gap between learning slopes for the congruent and incongruent conditions, as presented in Supplementary Figure 30.



<Supplementary Figure 30>

These findings suggest that models assuming a passive transition from affordance-driven behavior to goal-directed decision-making, irrespective of task context, are less likely to account for the observed faster adaptation in incongruent conditions compared to congruent ones for the same number of rewards.

We have clarified this in the revised Results section and added the following paragraph.

“To further rule out alternative explanations, we additionally implemented and tested a series of simpler models. These included two passive-decay models in which the affordance bias decreased over time according to either an exponential or hyperbolic decay function, as well as a restricted arbitration model in which the arbitration weight was determined solely by the reliability of the value-based decision system. All of these models were significantly outperformed by the full arbitration models that incorporate dynamic performance or reliability tracking across both systems (see Supplementary Fig. 29). Moreover, simulations from these simplified models failed to capture a key behavioral feature: the decreasing gap between congruent and incongruent conditions in the frequency of correct action selection as a function of the number of rewarded trials. In both the decay-based and the value-reliability-only models, this gap failed to decrease significantly (see Supplementary Fig. 30). These findings suggest that neither a passive transition away from affordance-driven behavior nor an arbitration process based only on value reliability is sufficient to explain the observed data. Rather, a dynamic, bidirectional arbitration mechanism—tracking the performance or reliability of both systems—is necessary to reproduce the behavioral effects.”

Details about the model implementation are also included in the Methods section, as follows:

“**Arbitration-weight-decay model.**
As a simple alternative to active dynamic arbitration models, we implement models with a decaying arbitration weight that gradually places greater weight on the value-based decision-making system. The model implementation is identical to that of the fixed arbitration model, except that the arbitration weight decays over the course of the task. In these models, the arbitration weight is a function of the trial index t. For example, in the case of exponential decay

$$w_{q,t} = w_q * w_d^t$$

where $0 < w_d \leq 1$; and in the case of hyperbolic decay,

$$w_{q,t} = w_q \frac{1}{1 + w_d \times t}$$

where 0<w_d. These models thus have 8 free parameters in total ($\beta, \alpha, b_{pinch}, b_{clench}, b_{poke}, k, w_q$, and w_d)”

These new models have also been added to the public codebase available at https://github.com/sanghyunyi/Affordance2025

Other comments:

-- The modelling endeavour is commendable, however the exposition of the modelling results can be better streamlined. For instance, comparing a Bayesian and RL model early on but then sticking with an RL formulation for the arbitration models is not well motivated/ explained.

Thank you for the thoughtful feedback. We agree that while we initially compared both Bayesian and RL frameworks, our transition to focusing exclusively on RL-based arbitration models in the later stages of model development may not have been sufficiently justified in the manuscript.

The Bayesian model offers a normative approach to decision-making by optimally integrating new learning signals. We explored several ways to incorporate affordance effects within this framework. One approach was to model the affordance effect as a prior, consistent with Bayesian principles. However, Bayesian models are generally less flexible when it comes to capturing suboptimal or non-normative learning behavior, particularly when participants do not update beliefs in an optimal manner. As a result, this approach placed natural limits on the model's ability to account for the behavioral data.

A second approach was to incorporate affordance as a bias and model adaptation as belief updates about the likelihood that a given action would lead to reward. Nonetheless, this class of models consistently yielded poorer fits to the behavioral data compared to even relatively simple RL models, which offered greater expressive power.

Given the flexibility of RL and its strong empirical grounding in modeling value-based learning, we ultimately chose to explore a broader set of RL-based model variants for arbitration.

We made these points clearer by adding the following sentences into the main section 2.3.

"We also tested the possibility that value learning could be supported by Bayesian inference rather than RL (Bayes model). Although Bayesian learning offers a normative framework for optimal belief updating, we found it less well-suited for capturing the flexible and sometimes suboptimal behavioral patterns observed in our task. We implemented several Bayesian variants, including models where affordance was incorporated either as a prior or as a bias influencing belief updates. However, these Bayesian models consistently showed poorer fits to the data compared to RL models (See Fig. 3) which allowed for more expressive adaptation to feedback. Given the greater flexibility of RL and its empirical success in modeling value-based choice behavior, we focused our subsequent modeling efforts on RL-based models."

-- It is not clear to me that the Affordance Prediction Error is a meaningful quantity. It combines incommensurable quantities: affordance compatibility and reward value. I also think that the notion of reliability is relatively fixed for affordance reliability. Reward reliability does increase as a function of learning, so I was wondering whether the authors have tried a simpler reliability model which resorts to the affordance bias only when value-based reliability is low (at the early trials).

We agree that the Affordance Prediction Error (APE) is a computational construct designed to capture deviations from expected affordance-based choices, and we acknowledge that it is not directly comparable to the Reward Prediction Error (RPE) in terms of scale or psychological interpretation. Our intention in introducing the APE was not to suggest a neural or conceptual equivalence with RPE, but rather to provide a tractable mechanism for monitoring the predictive performance, or reliability, of the affordance-based decision system within the broader arbitration framework. In this context, APE serves as a heuristic for estimating how effectively the affordance-based system guides behavior over time. Specifically, the system's reliability dynamically decreases when affordance-based choices fail to yield rewards and increases when affordance-compatible actions lead to successful outcomes.

We recognize that the concept of "reliability of affordance" can also be interpreted as the inherent strength of stimulus–action associations, which we captured using static affordance scores in our models. However, in response to the reviewer's comment, we additionally tested a simplified variant in which arbitration weights were determined solely by the reliability of the value-based system. This value-only reliability model performed worse than the full reliability-based arbitration model that tracks both systems (see Supplementary Figure 29). Moreover, this model failed to reproduce the key behavioral signature: a decreasing gap between congruent and incongruent conditions in correct action choices as a function of the number of rewarded trials (see Supplementary Figure 30).

We have clarified this in the revised section 2.3 and added the following sentences.

"To further rule out alternative explanations, we additionally implemented and tested a series of simpler models. These included two passive-decay models in which the affordance bias decreased over time according to either an exponential or hyperbolic decay function, as well as a restricted arbitration model in which the arbitration weight was determined solely by the reliability of the value-based decision system. All of these models were significantly outperformed by the full arbitration models that incorporate dynamic performance or reliability tracking across both systems (see Supplementary Fig. 29). Moreover, simulations from these simplified models failed to capture a key behavioral feature: the decreasing gap between congruent and incongruent conditions in the frequency of correct action selection as a function of the number of rewarded trials. In both the decay-based and the value-reliability-only models, this gap failed to decrease significantly (see Supplementary Fig. 30). These findings suggest that neither a passive transition away from affordance-driven behavior nor an arbitration process based only on value reliability is sufficient to explain the observed data. Rather, a dynamic, bidirectional arbitration mechanism—tracking the performance or reliability of both systems—is necessary to reproduce the behavioral effects."

Details about the model implementation are also included in the Methods section, specifically within the reliability arbitration model subsection, as follows:

"As a simpler alternative, we also tested a value-reliability arbitration model that relies solely on the reliability of the value-based decision-making system. This reduced model is identical to the full reliability-based model, except in the computation of the arbitration weight:

$$w_{q,t}\,(s) = \frac{1}{1 + exp(-\eta_q\, \chi_{q,t}(s) + \eta_0\ )}$$

This model thus has 9 free parameters."

This new model has also been added to the public codebase available at https://github.com/sanghyunyi/Affordance2025

-- The authors state that the RT behavioural effects are indistinguishable in early and late trials. However, the choice effects manifest differently at early trials (larger affordance bias) relative to late trials. It is not clear whether these two statements are at odds with each other.

We appreciate the reviewer's perspective and fully acknowledge that, in many decision-making frameworks such as the drift diffusion model (DDM), choice and RT are closely coupled, with longer RTs generally reflecting slower evidence accumulation due to weaker or more ambiguous information, or a greater distance between the starting point of the diffusion process and the decision boundary.

However, in the current paradigm, RT reflects a complex mixture of factors beyond decision uncertainty alone. These include trial progression, experimental condition, action type, reward status of the chosen action, and affordance compatibility. To disentangle these influences, we conducted mixed-effects regression analyses controlling for these variables, and found that the RT difference between affordance-compatible and incompatible actions remains robust across both early and late trials (see Supplementary Table 1).

Importantly, the persistence of this RT difference across learning suggests that it is not simply a function of changing decision certainty or learning-related dynamics. Instead, we interpret this RT effect as a motoric or perceptual facilitation consistent with stimulus–response compatibility, which exerts a stable influence irrespective of evolving reward contingencies. This is supported by the lack of significant change in the RT difference over time.

In contrast, the choice effect, reflected in a selection bias toward affordance-compatible actions, is dynamic and diminishes as participants learn which actions are more rewarding. This divergence between RT and choice effects suggests that the two measures are driven by distinct sources: RTs are shaped by stable sensorimotor mappings (i.e., affordance-based motor priming), while choices reflect the evolving arbitration between value-based and affordance-based decision systems.

Thus, while DDM frameworks predict a coupling between RT and choice under a unified evidence accumulation process, our findings suggest that, in this context, RT can be influenced by automatic motor biases that are largely independent of value-based decision-making. Therefore, the apparent dissociation between RT and choice effects in our task does not reflect a contradiction, but instead highlights the multifaceted nature of affordance influences on behavior, some of which are dynamic and decision-related, and others that are stable and rooted in sensorimotor priming.

To make these points clear, we added the following sentences to the discussion.

“Our findings also offer an important qualification to decision-making frameworks that assume a tight coupling between RT and choice, such as drift diffusion models. While these models predict that longer RTs reflect slower or more uncertain decisions along a unified evidence accumulation process, our results reveal a partial dissociation: the RT difference between affordance-compatible and incompatible actions remained robust across learning, whereas the choice bias diminished over time. This suggests that RTs in our paradigm are shaped by stable sensorimotor facilitation effects (e.g., stimulus–response compatibility), while choices reflect the dynamic arbitration between value-based and affordance-based systems. In this way, affordance exerts both automatic and adaptive influences on behavior, operating at distinct levels of the decision hierarchy.”

-- At the end of section 2.3 the authors acknowledge that their experimental paradigm does not allow them to delineate the temporal dynamics of processing value information and action affordances. I agree with their statements. Across these lines, I find that the slow vs. fast RT analyses is not informative. This analysis could be informative if we were sure that all slow responses are attributed to a higher decision boundary (with reference to an evidence accumulation framework). However, slow responses can also be attributed to other factors (low arousal levels or drift rate, longer non-decision time etc) which limits the inferential potential of a median split analysis in this paradigm.

We appreciate the reviewer's thoughtful observation and agree that our current paradigm is limited in its ability to precisely delineate the temporal dynamics of affordance and value-based processing. As correctly pointed out, interpreting slow versus fast RTs within an evidence accumulation framework would require strong assumptions about the sources of RT variability, such as drift rate, boundary separation, or non-decision time, that our task was not specifically designed to isolate or estimate. Therefore, we acknowledge that attributing slower RTs solely to affordance-stimulus compatibility is overly speculative, as alternative explanations like attentional fluctuations, or motor preparation differences cannot be ruled out.

Considering this, we recognize that the RT analysis offers limited inferential power and should be interpreted with caution. In the revised manuscript, we added the following paragraph to clarify that this analysis is exploratory in nature and does not provide conclusive insight into decision dynamics. We have also noted the importance of future work employing paradigms explicitly tailored to capture temporal aspects of decision-making, such as time-pressure manipulations or modeling approaches grounded in sequential sampling theory. We thank the reviewer for highlighting this important limitation and helping us more accurately frame our interpretations.

"Building on this, our findings highlight the importance of separating automatic motor influences from deliberative decision processes when interpreting RT patterns in value-based tasks. While RT is often used as a proxy for decision difficulty or uncertainty, our data suggest that this assumption may not hold in settings where stimulus–response compatibility introduces a persistent, low-level facilitation effect. To more clearly dissociate these processes, future work will benefit from paradigms that are explicitly designed to capture decision dynamics, such as those using time-pressure manipulations, sequential sampling models, or temporally resolved neural recordings. These approaches could help clarify how affordance-related signals interact with value-based evidence accumulation over time, and whether they modulate the decision process directly or operate through parallel, independent channels."

Reviewer #5 (Remarks to the Author):

This is an impressive multi-modal approach to investigating the interaction of action affordability and value-based decision making in a methodologically and theoretically strong paper. The authors admirably integrated the helpful comments by the reviewers in this revision, which has clarified several points about the strength of the modeling approach, the significance of the question, clarification on the methods and generalization of the paradigm/results. Helpful additions include supplementary figures 17-19, testing of new models (e.g., win-stay-lose-shift strategy) and addressing concerns about overparameterization. The authors also made their codebase publicly available to enhance transparency. Overall, this revised paper will make a significant contribution to the literature and is well-suited for publication in Nature Communication.

Thank you for the generous and encouraging feedback. We're pleased that the revisions, particularly the additional analyses, model comparisons, and shared code, helped clarify and strengthen the manuscript. We appreciate your thoughtful engagement throughout the review process and are grateful for your support in recommending the paper for publication.

Response to reviewer #4

Thank you for your positive feedback. We appreciate your careful review and are glad that our revisions have addressed all of your concerns