Archival Report

Biological Psychiatry: CNNI

Computationally Informed Insights Into Anhedonia and Treatment by Kappa Opioid Receptor Antagonism

Bilal A. Bari, Andrew D. Krystal, Diego A. Pizzagalli, and Samuel J. Gershman

ABSTRACT

BACKGROUND: Anhedonia, the loss of pleasure, is prevalent and impairing. Parsing its computational basis promises to explain its transdiagnostic character. One manifestation of anhedonia, reward insensitivity, may be linked to limited memory. Furthermore, the need to economize on limited memory engenders a perseverative bias toward frequently chosen actions. Anhedonia may also be linked with deviations from optimal perseveration for a given memory capacity, a pattern that causes inefficiency because it results in less reward for the same memory cost.

METHODS: To test these hypotheses, we applied a theory of optimal decision making under memory constraints that decomposes behavior into a memory component and an efficiency component. We applied this theory to behavior on the Probabilistic Reward Task, a reward learning paradigm that has been validated in anhedonia, and performed secondary analysis of a randomized controlled trial testing kappa opioid receptor (KOR) antagonism for anhedonia (n = 24 KOR; n = 31 placebo), as well as analyses of 3 other datasets (n = 100, 66, 24, respectively). We fit a resource-bounded reinforcement learning model to behavior.

RESULTS: Across clinical and nonclinical populations, anhedonia was associated with deficits in efficiency but not memory. The reinforcement learning models demonstrated that deficits in efficiency arise from the inability to perseverate optimally. KOR antagonism, which likely elevates tonic dopamine, increases both memory and efficiency, and the model demonstrated that this arises from increased reward sensitivity and perseveration.

CONCLUSIONS: Therefore, KOR antagonism has distinct cognitive effects, only one related to anhedonia. These findings have potential implications for the applications of KOR antagonists.

https://doi.org/10.1016/j.bpsc.2025.05.011

Anhedonia, the loss of pleasure or lack of reactivity to pleasurable stimuli, is observed in many psychiatric disorders (1–9), suggesting a common mechanism across disorders. The most systematic attempts to formalize this common mechanism have utilized concepts from reinforcement learning (10). Early models posited that anhedonia corresponds to a reduction in reward sensitivity (11,12), but the predictions of these models have not been consistently validated, suggesting a more complex picture (13). Here, we argue that one neglected source of complexity is the interplay between reward sensitivity and cognitive capacity limits.

In reinforcement learning theory, states (e.g., stimuli, context) are mapped to actions by a learned policy. The amount of memory needed to store a policy is dictated by the mutual information between states and actions; any physical system (such as the brain) has a limited memory capacity. One implication of limited capacity is reward insensitivity, and thus some aspects of anhedonia may arise from cognitive resource limitations.

Under capacity limits, policies must be compressed by discarding some state information (14–16). This results in the tendency to reuse frequently chosen actions across multiple states—a form of perseveration, the tendency to repeat actions independently of their reinforcement history. The theory of policy compression is normative: It specifies an optimal level of perseveration for a given capacity limit. Empirically, compression strategies may differ, with some policies yielding more reward than others for the same cost. We refer to deviations from optimal perseveration as inefficiency because it results in a suboptimal use of finite memory (less reward for the same memory utilization). This phenotype is conceptually distinct from capacity and can be measured separately. We argue here that capacity and efficiency may be key phenotypes for understanding cognitive disturbances in anhedonia. We show that these can be estimated from behavioral data on a widely used behavioral assay, the Probabilistic Reward Task (PRT), and that they reveal new aspects of anhedonia that otherwise would have been invisible.

We also address the underlying neural mechanisms and treatment implications. Our previous work suggested that tonic dopamine should determine the allocation of cognitive resources for task performance based on reward rate (17,18). Therefore, reduction in tonic dopamine should produce insensitivity of task performance to reward rate (19). It stands

© 2025 Society of Biological Psychiatry. Published by Elsevier Inc. All rights reserved, including those for text and data mining, AI training, and similar technologies.

ISSN: 2451-9022

to reason that increasing tonic dopamine should increase reward sensitivity. We demonstrate that this is consistent with the effects of kappa opioid receptor (KOR) antagonism, which elevates tonic dopamine (20–24). We found that efficiency also increases, suggesting that tonic dopamine may not only determine the amount of resources available but also the efficiency of their allocation. Mechanistically, this might be implemented through dopamine-dependent changes in learning rate for perseveration. Finally, we found that anhedonia was associated with changes in efficiency but not memory, highlighting the clinical utility of distinguishing these computational phenotypes.

METHODS AND MATERIALS

Probabilistic Reward Task

The PRT is a reward-based learning task that has been validated in anhedonia (25,26). On each trial, participants are presented 1 of 2 perceptually similar stimuli and asked to report the stimulus that they perceived. Participants completed either 200 trials (KOR dataset) or 300 trials (all other datasets). The critical feature of the task was that correct responses for one stimulus yielded reward with probability 60%, and correct responses for the other stimulus yielded reward with probability 20% (stimuli/responses counterbalanced across participants).

Policy Compression: A Capacity Limit on Decisions

Policy compression is an application of rate-distortion theory, a subdiscipline of information theory, that treats decision making as communication across a capacity-constrained channel (14,16). Under policy compression, state information (stimuli, in our case) is transmitted across a channel to produce a policy, $\pi(a|s)$, a probability distribution of actions conditioned on state. We assume that participants cannot perfectly transmit information across this channel but must keep it below a capacity constraint. This requires that participants trade-off the utility of their decisions with the cognitive costs, which we define as the mutual information between states and actions, or the policy complexity. In general, policies that result in higher reward incur a greater memory cost. We treat policy complexity as the memory applied to the task and the difference between optimal reward and empirical reward, for a given policy complexity, as the efficiency of memory usage.

Reinforcement Learning Modeling

We constructed the following resource-bounded *Q*-learning model, motivated by policy compression. It estimates action values, $Q(s,a)$, and the marginal action probability, $P(a)$, to generate a policy and contains 3 parameters (α_{learn}, α_{persev}, and β):

$$\Delta Q(s,a) = \alpha_{\text{learn}}[r - Q(s,a)]$$

$$\begin{aligned}\Delta P(a) &= \alpha_{\text{persev}}[\pi(a|s) - P(a)] \\ \pi(a|s) &\propto exp[(\beta Q(s,a) + log(P(a)))]\end{aligned} \quad (1)$$

where $r = 1$ if the current trial is rewarded and 0 otherwise. This model contains a mechanism to allow KOR treatment to multiplicatively scale α_{persev} and β in the following manner:

$$\begin{aligned}\alpha_{\text{persev}} &= \alpha_{\text{persev,baseline}} \cdot 10^{s_{\text{persev,treatment}}} \\ \beta &= \beta_{\text{baseline}} \cdot 10^{s_{\text{beta,treatment}}}\end{aligned} \quad (2)$$

For the online study, we scaled parameters as a function of the *z*-scored anhedonia inventory scores (using the Snaith-Hamilton Pleasure Scale [SHAPS]) in the following manner:

$$\begin{aligned}\alpha_{\text{persev}} &= \alpha_{\text{persev,baseline}} \cdot 10^{s_{\text{persev}} \cdot \text{SHAPS}} \\ \beta &= \beta_{\text{baseline}} \cdot 10^{s_{\text{beta}} \cdot \text{SHAPS}}\end{aligned} \quad (3)$$

We initialized $Q(s, a)$ at 0 and $P(a)$ at 0.5, and we assumed scaling terms equaled 0 on sessions without treatment. We included all trials for analysis. Models were fit with the probabilistic programming language Stan.

RESULTS

Policy Complexity and Efficiency in Anhedonia After KOR Antagonism

We performed a secondary analysis of an 8-week, multicenter, placebo-controlled, double-blind randomized trial to test the effects of KOR antagonism on anhedonia (Figure 1A) (27,28). Because this trial identified a significant treatment effect of KOR antagonism for anhedonia (as measured by SHAPS), we sought to understand the cognitive basis of this improvement. We analyzed a total of 55 participants (KOR antagonist group: $n = 24$; placebo group: $n = 31$) who completed both baseline and posttreatment PRT. Owing to previously reported baseline differences in anhedonia between the 2 groups (mean SHAPS ± SD: placebo 33.03 ± 5.54; KOR 37.29 ± 8.89; $p = .0338$), we analyzed the pretreatment groups separately.

The PRT is a reward-based decision-making task that asks participants to discriminate 2 similar stimuli (25,26) (Figure 1B). Unbeknownst to participants, one of the 2 stimuli yields reward more often than the other when correctly identified. According to the theory of policy compression (16), performance in this task (average reward) depends on the amount of information that participants encode about the underlying state (i.e., the stimulus identity), quantified by the mutual information between states and actions—a participant's policy complexity. Each participant is assumed to have a capacity limit (upper bound on policy complexity), which delimits their achievable performance. If participants maximally utilize their capacity, their average reward should fall along an optimal reward-complexity frontier, as shown in Figure 2A, B. In the PRT, maximal reward can be obtained at a policy complexity of 1 bit, corresponding to a policy that perfectly discriminates the 2 stimuli. At the other extreme, a participant with no capacity will generate a policy that ignores the stimuli entirely. Participant policies tend to lie close to the optimal frontier, indicating that they are utilizing most of their capacity. At the low end of the policy complexity range, participant policies fall off the optimal frontier (Figure 2F, G), indicating underutilization of resources (inefficiency), a pattern that has also been observed in previous studies (15,29). Consistent with theory, participants with lower policy complexity tended to perseverate more (effect of policy

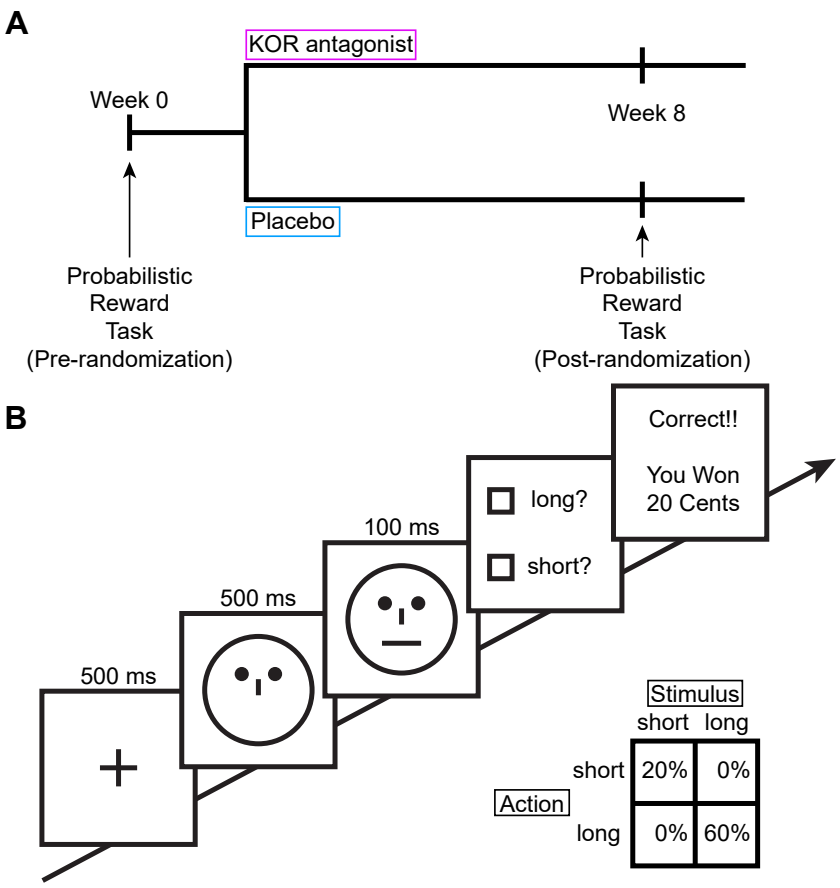

Figure 1. Trial and task design. **(A)** Participants were randomized to 8 weeks of placebo (n = 31) or a kappa opioid receptor (KOR) antagonist (n = 24) and completed the Probabilistic Reward Task (PRT) at baseline and at week 8. **(B)** On each trial of the PRT, participants fixated on a cross, followed by the presentation of a face without a mouth, followed by either a short (11.5 mm) or long (13 mm) mouth in the face. Participants responded by pressing one of the 2 keyboard keys and completed 200 trials in 2 blocks of 100 trials. The bottom right shows an example reward schedule where the long stimulus is rewarded more often than the short stimulus. The mapping between response, stimulus, and reward was counterbalanced between participants.

complexity on probability of choosing the richer option, coefficient = −0.251, p = .0188). See Figure S1 for further intuition into the relationship between policy complexity and behavior.

At 8 weeks, placebo treatment resulted in a decrease in both policy complexity and reward, while KOR antagonism yielded an increase in both (Figure 2C). This resulted in significant between-group differences for both policy complexity (mean change in policy complexity [posttreatment minus baseline] ± SEM: placebo, −0.0245 ± 0.0141; KOR, 0.0281 ± 0.0211; p = .0362) (Figure 2D) and reward (mean change in reward [posttreatment minus baseline] ± SEM: placebo, $-0.0165 \pm 5.61 \times 10^{-3}$; KOR, $0.0154 \pm 6.53 \times 10^{-3}$; $p = 4.81 \times 10^{-4}$) (Figure 2E). Following treatment, the KOR group also became significantly more efficient compared with the placebo group (mean change in inefficiency [posttreatment minus baseline] ± SEM: placebo, $0.0130 \pm 4.80 \times 10^{-3}$; KOR, $-0.0109 \pm 4.04 \times 10^{-3}$, $p = 5.68 \times 10^{-4}$) (Figure 2H). Increased efficiency can be interpreted as a tendency to more optimally select the richer option (Figure S1B). Consistent with this interpretation, we found that participants were more optimal in selecting the richer option after KOR treatment (posttreatment minus baseline ± SEM: placebo, 0.0502 ± 0.0221; KOR, -0.0448 ± 0.0196, $p = 2.96 \times 10^{-3}$) (Figure S2). Thus, KOR antagonism increased average reward through a combination of increasing both policy complexity and efficiency.

Policy compression makes the additional prediction that more complex policies should result in slower response times because the brain must inspect more bits to find a coded state (16,18,30). We found that KOR antagonism, relative to placebo, slowed participants down (mean change in response times [posttreatment minus baseline] ± SEM: placebo, −59.3 ± 23.4 ms; KOR, 13.6 ± 20.4 ms; p = .0274).

To better understand how KOR treatment changed the relationship between inefficiency and policy complexity, we fit a linear mixed-effects model predicting inefficiency as a function of policy complexity, treatment, and time. We identified 2 relevant effects: a significant treatment × time interaction (coefficient = −0.0405, $p = 4.23 \times 10^{-5}$), which has the effect of lowering the intercept, and a significant policy complexity × treatment × time interaction (coefficient = 0.187, $p = 1.76 \times 10^{-3}$), which has the effect of increasing the slope. The combination of the change in intercept and slope has the net effect of increasing efficiency as a function of policy complexity, revealing that KOR treatment increases efficiency independent of its changes to complexity. We will develop this insight further with our reinforcement learning modeling. Overall, these results suggest 2 orthogonal effects

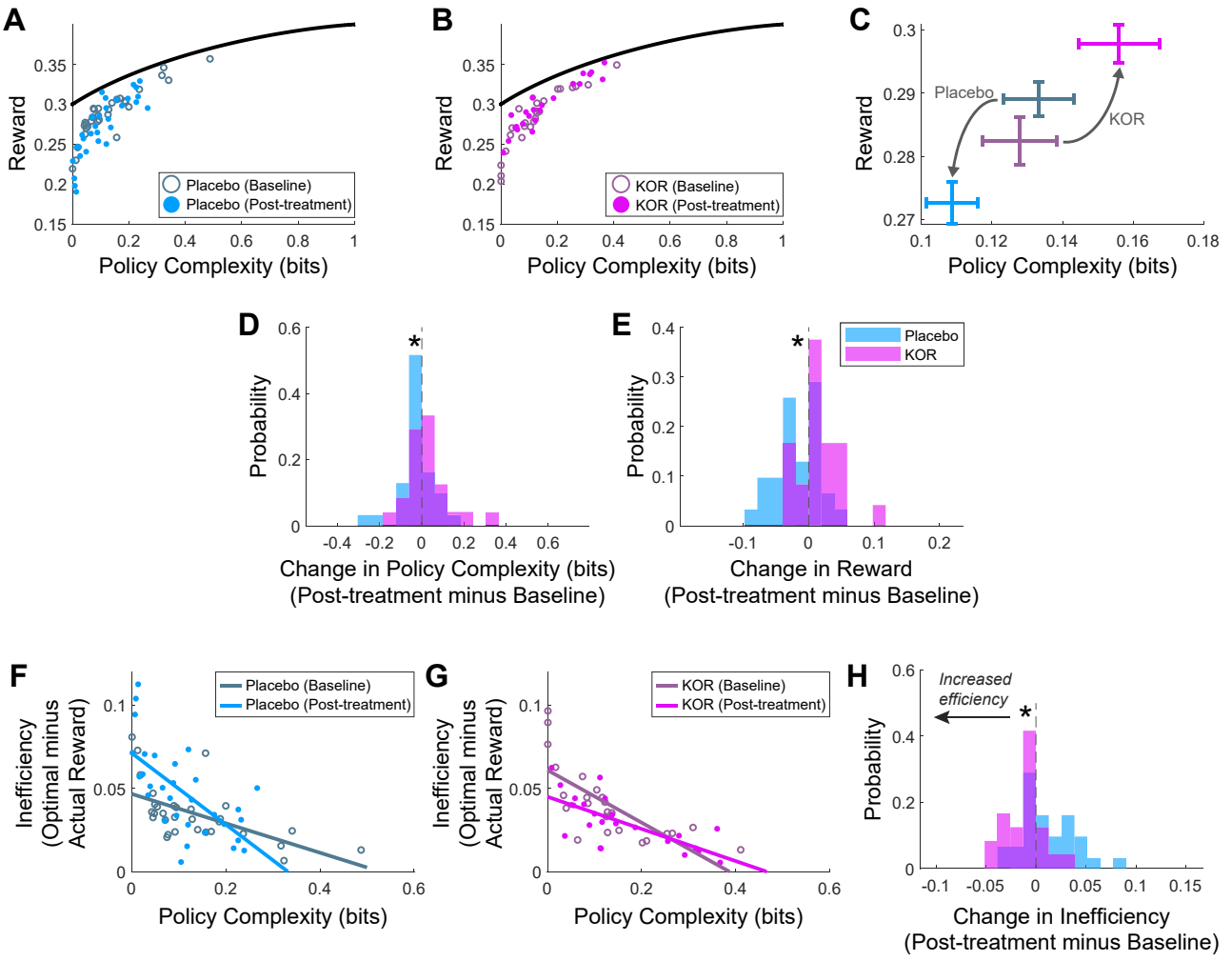

Figure 2. Changes in policy complexity and efficiency as a function of kappa opioid receptor (KOR) antagonism. **(A, B)** Reward-complexity relationship for the placebo and KOR groups at baseline and posttreatment. The black line shows the reward-complexity frontier, which indicates the optimal reward as a function of policy complexity. **(C)** Mean ± SEM reward-complexity relationship as a function of treatment (placebo or KOR antagonism) and time (baseline or posttreatment). **(D)** Change in policy complexity (posttreatment minus baseline) as a function of treatment. **(E)** Change in reward (posttreatment minus baseline) as a function of treatment. **(F, G)** Relationship between inefficiency and complexity for the placebo and KOR groups. Overlaid lines are from a linear mixed-effects model fitting inefficiency as a function of policy complexity, treatment, and time. **(H)** Change in inefficiency (posttreatment minus baseline) as a function of treatment. *$p < .05$.

of KOR treatment: increases in complexity and increases in efficiency. Stated another way, participants gain increased cognitive resources and make better use of those resources.

Reinforcement Learning Model of KOR Antagonism

We developed a cost-sensitive reinforcement learning model to gain insight into how KOR antagonism affects decision making. We adapted a *Q*-learning model, which is ubiquitous in the reinforcement learning literature (31). This model estimates the expected reward associated with each action for each stimulus (called *Q*-values) and updates these estimates by learning from the outcome (presence or absence of reward). Because the optimal policy under policy compression contains a marginal action probability term to engender perseveration (state-independent actions), we augmented our model with a marginal action probability term that was similarly estimated on a trial-by-trial basis. Our model contained a reward learning rate, α_{learn}, to govern the learning of action values; a perseveration learning rate, α_{persev}, to govern the learning of the marginal action probability; and a reward sensitivity parameter, β, that determines the balance between action values and perseveration in driving behavior. The β parameter is linked to capacity, where higher capacity is associated with higher values of β. Given the structure of our model, β is equivalent to a parameter scaling reward magnitude, as has been posited in anhedonia (12).

To model the effects of treatment, we allowed KOR and placebo to scale these parameters. Based on formal model comparison (Table S1), we selected a model that separately

scaled the perseveration learning rate, α_{persev}, and the reward sensitivity, β, as a function of treatment. We confirmed that our model could recover α_{persev} and β, the parameters of interest (Table S2). To provide confidence in the ability of our model to capture key characteristics of the data, we first fit the model to participant data and then had the model perform the PRT (using the parameter estimates for each participant) to generate a synthetic dataset (Figure S8). This simulated dataset captured all key features of our data (see the Supplement).

Having confirmed that our model could generate realistic data and recover parameters of interest, we turned our attention to parameter estimates to better understand how treatment affected decision making. We found that placebo and KOR antagonism scaled the perseveration learning rate, α_{persev}, in opposite directions (posterior 95% credible interval; placebo, −2.96 to −0.82; KOR, 0.61 to 1.96) (Figure 3A). The difference between KOR antagonism and placebo corresponds to the net effect of treatment on α_{persev}, which was positive and excluded 0, showing that treatment increased perseveration (difference in posterior 95% credible interval [KOR minus placebo], 1.77 to 4.32) (Figure 3B). We similarly found that placebo and KOR antagonism scaled the reward sensitivity, β, in opposite directions (posterior 95% credible interval; placebo, −0.143 to −0.050; KOR, 0.037 to 0.138) (Figure 3C), with a treatment effect that was positive and excluded 0 (difference in posterior 95% credible interval [KOR minus placebo], 0.114 to 0.254) (Figure 3D).

To gain insight into how scaling these parameters affects decision making, we simulated datasets where we only changed parameters of interest (Figure S3 and Table S3). Increasing only α_{persev} produces an increase in efficiency and a small decrease in policy complexity. The increase in efficiency manifests as a change in the intercept, but not the slope, of the relationship between inefficiency and policy complexity. Increasing only β produces a relatively large increase in policy complexity, which is consistent with the theoretical link between larger β and increased capacity. It also produces an increase in efficiency for low-complexity policies. Increasing both α_{persev} and β, like we find for KOR antagonism, produces both an increase in policy complexity and an increase in efficiency. The increase in efficiency manifests as a change in both the intercept (decrease) and the slope (increase) of the relationship between inefficiency and policy complexity, like our empirical findings.

We gained insight into the relationship between KOR antagonism and optimal behavior by visualizing the relationship between α_{persev}, β, and reward, while holding α_{learn} fixed (Figure 3E). As β increases, for the optimal α_{persev}, the net reward obtainable also increases, consistent with our theory linking higher β to higher capacity and higher capacity to greater reward. We also find that increasing perseverative learning is most beneficial at lower values of β (i.e., lower capacity), consistent with the idea that perseveration is increasingly optimal as participants become more resource limited. In Figure 3F, we can see that the effect of KOR antagonism is to shift both α_{persev} and β closer to an optimal regime. A notable finding is the increased α_{persev} at baseline for the placebo group relative to the KOR group. This is consistent with the baseline difference in SHAPS between these groups, with the placebo group having lower SHAPS: The larger α_{persev} estimates for this group are closer to the optimal regime and are consistent with less severe anhedonia.

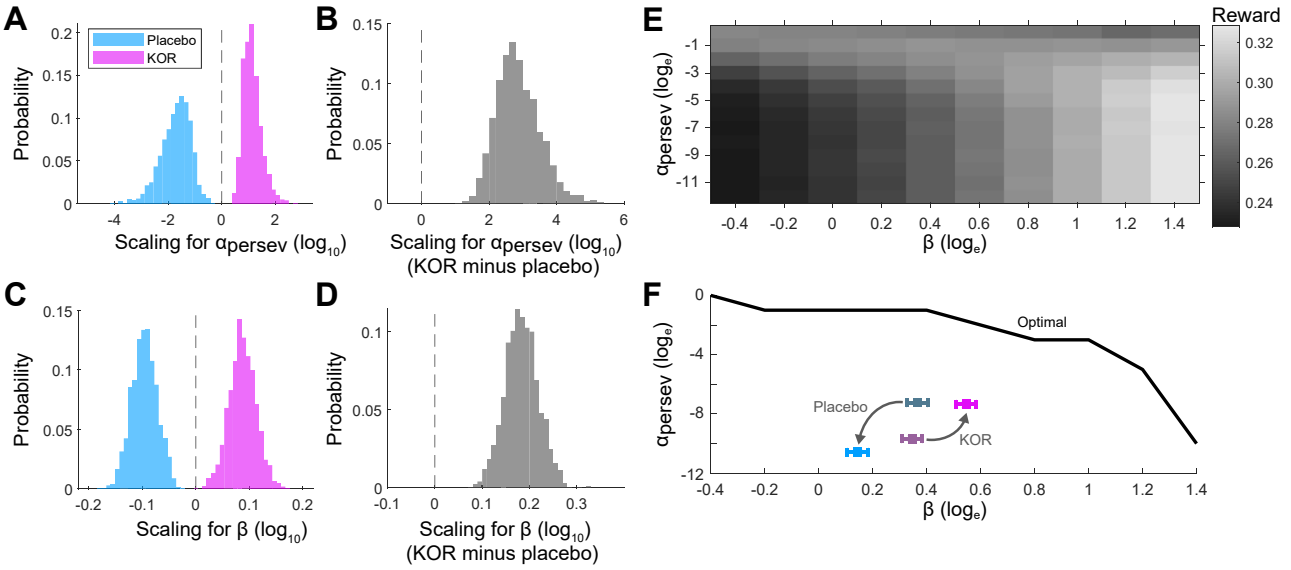

Figure 3. Scaling of reinforcement learning parameters as a function of kappa opioid receptor (KOR) antagonism. **(A)** Posterior distribution of parameter values for scaling of α_{persev} as a function of treatment. Scaling is multiplicative, where values > 0 indicate that treatment increases the parameter value, whereas values < 0 indicate that treatment decreases the parameter value. **(B)** Posterior distribution of treatment effect for scaling α_{persev}, estimated as the difference in scaling between KOR and placebo. **(C)** Posterior distribution of parameter values for scaling of β as a function of treatment. **(D)** Posterior distribution of treatment effect for scaling β. **(E)** Heatmap showing mean reward obtained as a function of α_{persev} and β. **(F)** Effect of treatment in parameter space. The black line shows the optimal α_{persev} for each value of β.

Policy Complexity and Efficiency as a Function of Hedonic Tone

Because the original study identified a significant improvement in SHAPS following KOR antagonism (27), we sought to identify which mechanism—increased policy complexity, increased efficiency, or both—is associated with anhedonia. We first examined the relationship between hedonic tone and reward learning in a nonclinical population. We recruited 100 participants from Amazon Mechanical Turk and implemented a version of the PRT suitable for online delivery (32). Participants completed SHAPS and reported a wide range of scores (mean SHAPS ± SD: 11.45 ± 6.54, range 0–36). We show the reward-complexity relationship in Figure 4A. For visualization purposes only, we performed a median split of participants on the basis of SHAPS.

Unlike the effects of KOR antagonism, we found that SHAPS did not predict policy complexity (coefficient = -5.24×10^{-3}, $p = .241$). However, we did identify a significant relationship with inefficiency. We fit a linear regression predicting inefficiency as a function of SHAPS and policy

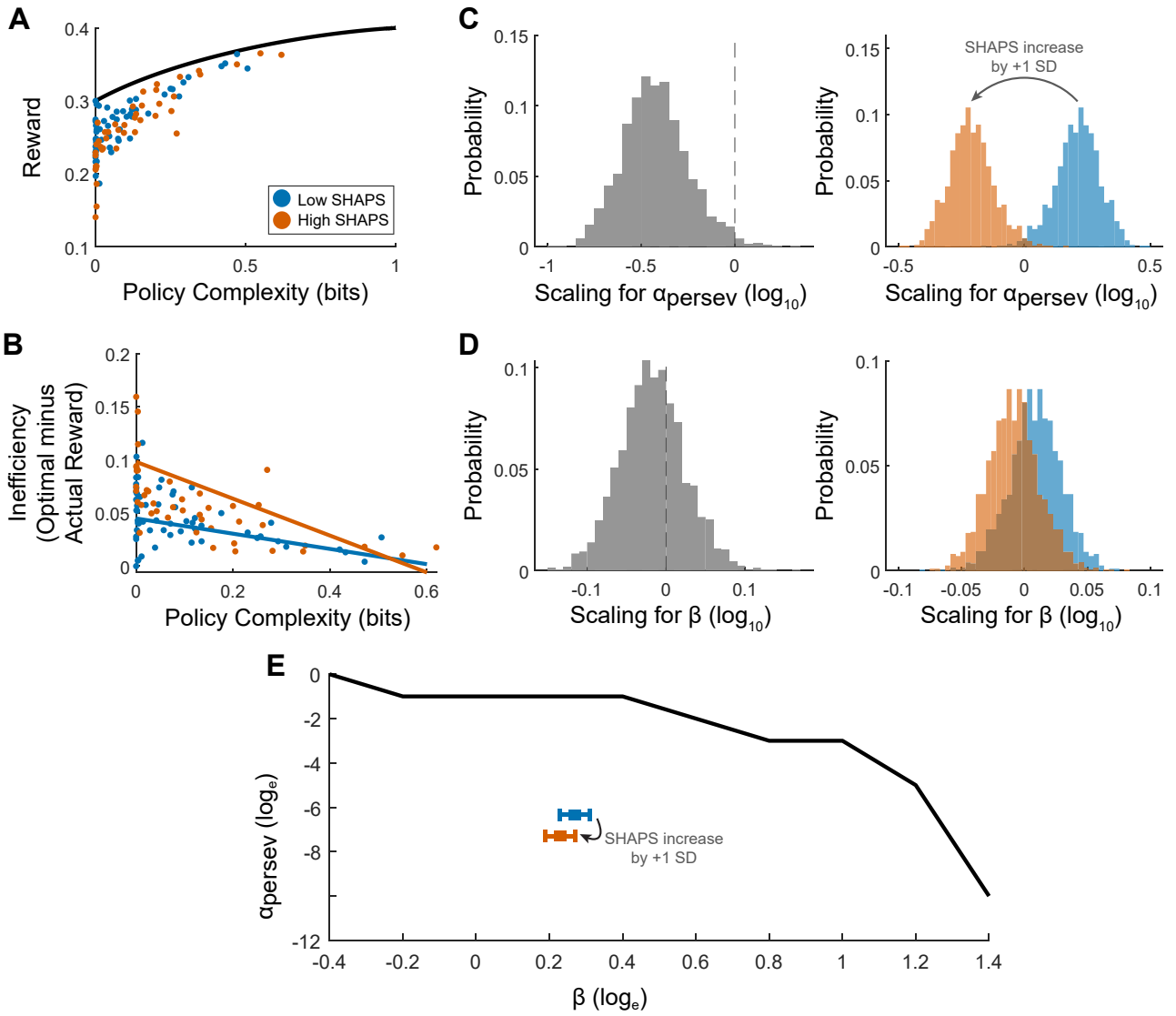

Figure 4. Changes in complexity, efficiency, and reinforcement learning parameters as a function of hedonic tone. **(A)** Reward-complexity trade-off as a function of hedonic tone. For illustration only, participants are median split on the basis of Snaith-Hamilton Pleasure Scale (SHAPS) scores into low SHAPS (low anhedonia) and high SHAPS (high anhedonia). **(B)** The inefficiency-complexity relationship as a function of hedonic tone. For illustration only, the color lines are regression fits denoting extremes of SHAPS in our dataset (blue is lowest SHAPS = 0; orange is highest SHAPS = 36). **(C)** Left: posterior distribution of parameter values for scaling of α_{persev} as a function of SHAPS. Right: an example demonstrating scaling for an increase in SHAPS of 1 SD (from 0.5 SD below the mean [blue] to 0.5 SD above the mean [orange]). **(D)** Left: posterior distribution of parameter values for scaling of β as a function of SHAPS. Right: scaling for the same increase in SHAPS. **(E)** Effect of variation in SHAPS in parameter space. The black line shows the optimal α_{persev} for each value of β.

complexity and identified a significant intercept change (coefficient for effect of SHAPS = 9.55×10^{-3}, $p = 6.54 \times 10^{-3}$) but not a significant slope change (coefficient for SHAPS × policy complexity interaction = −0.0182, $p = .394$). Given our simulations exploring the effects of changing parameters (Figure S3), a change of intercept without a change of slope is consistent with hedonic tone affecting perseveration (α_{persev}) and not capacity (β).

We reanalyzed two previous PRT datasets and found similar effects on the relationship between inefficiency and policy complexity (Figure S6). The first was a transdiagnostic sample of patients (control group: $n = 25$; clinical group: $n = 41$, 18 with bipolar disorder, 23 with major depressive disorder) (33,34). These groups differed significantly in baseline anhedonia (mean anhedonic Beck Depression Inventory-II subscore ± SD: control, 0.72 ± 1.02; clinical, 5.22 ± 3.78, $p = 2.22 \times 10^{-7}$; mean Mood and Anxiety Symptom Questionnaire anhedonic depression subscale ± SD: control, 51.5 ± 12.6; clinical, 77.1 ± 19.3, $p = 1.45 \times 10^{-7}$). Consistent with differences in anhedonia, when we analyzed inefficiency as a function of policy complexity and group, we identified a significant intercept difference (coefficient for clinical group = 5.29×10^{-3}, $p = .022$) without a concurrent slope difference (coefficient for policy complexity × clinical group interaction = -1.71×10^{-3}, $p = .443$). We also found no difference in policy complexity between the 2 groups (mean policy complexity ± SEM: control, 0.371 ± 0.043; clinical, 0.333 ± 0.024; $p = .412$).

The second dataset that we analyzed tested the hypothesis that decreased dopamine impairs reward learning and leads to anhedonia (35,36). In this double-blind study, participants received either placebo or low-dose pramipexole, thought to reduce phasic dopamine release, and performed the PRT (placebo group: 13; pramipexole group: 11) (37). The original study found that pramipexole administration led to impaired reward learning. When we analyzed inefficiency as a function of policy complexity and treatment, we identified a significant intercept effect (coefficient for treatment = 7.82×10^{-3}, $p = .043$) without a significant slope effect (coefficient for policy complexity × treatment = -1.90×10^{-3}, $p = .615$). We also found no difference in policy complexity as a function of treatment (mean policy complexity: placebo, 0.297 ± 0.043; pramipexole, 0.319 ± 0.057; $p = .757$).

Reinforcement Learning Model of Hedonic Tone

Next, we fit a reinforcement learning model to the nonclinical population dataset. This model was similar to the one that we used for the KOR dataset, except now we allowed α_{persev} and β to scale as a function of SHAPS. We found that increases in SHAPS were associated with less perseveration (posterior 95% credible interval, −0.739 to −0.046) (Figure 4C). In contrast, anhedonia had no effect on modulating β, in contrast to KOR antagonism (posterior 95% credible interval, −0.097 to 0.066) (Figure 4D). In parameter space, the net effect of an increase in SHAPS is to move participants away from an optimal regime (Figure 4E). Taken together, these data support the notion that hedonic tone spans the axis of efficiency, not capacity.

DISCUSSION

We leveraged a theory of resource-limited reinforcement learning to shed light on the cognitive structure of anhedonia. Building on previous work demonstrating impairments in reward sensitivity, we decomposed these impairments into separate effects of policy complexity (state dependence of an action policy) and efficiency (utilization of cognitive resources). We found that KOR antagonism affected both of these measures, whereas anhedonia was associated only with reduced efficiency.

The finding that anhedonia was not associated with reduced complexity is surprising, in part because complexity determines reward sensitivity, and reward insensitivity appears to be the cardinal feature of anhedonia [but see (13) for more nuance]. There are several explanations for this apparent disconnect. One is that the subjective experience of anhedonia may be more related to the psychological concept of liking, the pleasure associated with reward, rather than wanting, the motivation furnished by reward learning (38), although deficits in both liking and wanting play a role in anhedonia (39) [but see (40,41) for challenges in measuring liking]. In our paradigm, reward sensitivity is related to wanting, which would render the PRT an inappropriate assay to measure deficits in liking. Furthermore, SHAPS is not designed to disambiguate these different aspects of reward processing, but newer scales such as the Dimensional Anhedonia Rating Scale (42), the Temporal Experience of Pleasure Scale (43), and the Positive Valence Systems Scale (44) provide insight into the multidimensional nature of anhedonia. It should be stated that the PRT's success in capturing deficits in reward learning in anhedonia strongly suggests that deficits in wanting are critical to understanding anhedonia, even if the link between the subjective experience of anhedonia and wanting is not immediate.

Our finding that anhedonia was associated with reduced efficiency has potential clinical relevance. Under our computational framework, perseveration is closely related to habits, because habits can be similarly thought of as state-independent actions within a particular context (45,46). A prediction of our findings is that anhedonia may not only manifest as a deficit in perseveration but may also manifest as a deficit in habit formation. Intriguingly, recent work on the origin of habits has revealed that they are largely divorced from reward (47). If true, this would highlight a cognitive deficit in anhedonia unrelated to reward processing. Together, our findings motivate a future research program that studies habit formation in anhedonia, which is important both for better understanding this symptom and because it may form the basis of clinically relevant behavioral interventions.

A key limitation is that our application of policy compression assumes no perceptual ambiguity. This assumption is imperfect because, by design, the PRT introduces perceptual ambiguity with a combination of perceptually similar stimuli and short stimulus durations. However, perceptual ambiguity does not fully capture behavior on the PRT (Figure S7). Participants consistently perform better than a policy that assumes perfect discrimination, subject to perceptual ambiguity. This difference is because participants gravitate toward the more rewarding response, a bias that arises during decision making. The bias accounts for the difference between policy compression and

perceptual ambiguity curves, which are equal only at perfect performance. This bias is normatively justified under policy compression, which therefore provides the computational language needed to understand behavior, both in health and anhedonia. That said, the most likely descriptor of behavior is a combination of perceptual ambiguity and policy compression. One reason that policy compression remains relevant is that information transmission across synapses is metabolically costly, incentivizing the brain to limit information transmission (48). Note that this cost is also related to perceptual performance, which is subject to similar information-theoretical costs (17,49). Future studies should parameterize perceptual ambiguity to better resolve its contribution to behavior so that the exact contribution of policy compression can be better isolated.

Conclusions

We leveraged computational principles to identify 2 mechanisms of action of KOR antagonism—one related to anhedonia (increase in efficiency) and one unrelated to anhedonia (increase in policy complexity). We hypothesize that the increase in complexity can be leveraged for other indications, including possibly cognitive deficits in psychosis. Our results provide a clear example of the potential for computational psychiatry to provide transdiagnostic insights that integrate across levels of analysis.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institute of Mental Health (Grant Nos. R25MH094612 [to BAB], R37MH068376 and R01MH101521 [to DAP], and HHS-N271-2012-000006-I [to ADK]) and by the Harvard Brain Science Initiative Bipolar Disorder Seed Grant (to SJG).

BAB and SJG developed the theoretical framework. BAB, ADK, DAP, and SJG designed research and collected data. BAB analyzed the data. All authors wrote the article. All authors read and approved the final version of the article.

A previous version of this article was published as a preprint on medRxiv: https://doi.org/10.1101/2024.04.09.24304873.

ADK has been a consultant for Eisai, Axsome, Big Health, Harmony, Idorsia, Jazz, Janssen, Takeda, Millenium Merck, Neurocrine, Neurawell, Otsuka, Evecxia, and Sage Research and has received support from the National Institutes of Health, the Ray and Dagmar Dolby Family Fund, Janssen, Jazz, Neurocrine, Attune, Harmony, and Axsome. Over the past 3 years, DAP has received consulting fees from Boehringer Ingelheim, Compass Pathways, Engrail Therapeutics, Karla Therapeutics, Neumora Therapeutics, Neurocrine Biosciences, Neuroscience Software, Otsuka, Sage Therapeutics, Sama Therapeutics, Sunovion Therapeutics, and Takeda; he has received honoraria from the American Psychological Association, Psychonomic Society, and Springer (for editorial work), as well as Alkermes; he has received research funding from the Brain and Behavior Research Foundation, Dana Foundation, Wellcome Leap, Millennium Pharmaceuticals, and NIMH; he has received stock options from Compass Pathways, Engrail Therapeutics, Neumora Therapeutics, and Neuroscience Software. DAP has a financial interest in Neumora Therapeutics, which has licensed the copyright to the PRT through Harvard University. The interests of DAP were reviewed and are managed by McLean Hospital and Mass General Brigham in accordance with their conflict-of-interest policies. No funding from these entities was used to support the current work, and all views expressed are solely those of the authors. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts (BAB); McLean Hospital, Harvard Medical School, Belmont, Massachusetts (BAB, DAP); Department of Psychiatry, University of California San Francisco, San Francisco, California (ADK); Departments of Psychiatry and Behavioral Sciences and Radiology, Duke University School of Medicine, Durham, North Carolina (ADK); Department of Psychology and Center for Brain Science, Harvard University, Cambridge, Massachusetts (SJG); and Center for Brains, Minds, and Machines, Massachusetts Institute of Technology, Cambridge, Massachusetts (SJG).

DAP and SJG are joint senior authors.

Address correspondence to Bilal A. Bari, M.D., Ph.D., at bbari@mgh.harvard.edu.

Received Feb 24, 2025; revised May 8, 2025; accepted May 18, 2025.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2025.05.011.

REFERENCES

1. Gard DE, Kring AM, Gard MG, Horan WP, Green MF (2007): Anhedonia in schizophrenia: Distinctions between anticipatory and consummatory pleasure. Schizophr Res 93:253–260.
2. Kashdan TB, Zvolensky MJ, McLeish AC (2008): Anxiety sensitivity and affect regulatory strategies: Individual and interactive risk factors for anxiety-related symptoms. J Anxiety Disord 22:429–440.
3. Hatzigiakoumis DS, Martinotti G, Giannantonio MD, Janiri L (2011): Anhedonia and substance dependence: Clinical correlates and treatment options. Front Psychiatry 2:10.
4. Chevallier C, Grèzes J, Molesworth C, Berthoz S, Happé F (2012): Brief report: Selective social anhedonia in high functioning autism. J Autism Dev Disord 42:1504–1509.
5. Meinzer MC, Pettit JW, Leventhal AM, Hill RM (2012): Explaining the covariance between attention-deficit hyperactivity disorder symptoms and depressive symptoms: The role of hedonic responsivity. J Clin Psychol 68:1111–1121.
6. Nawijn L, van Zuiden M, Frijling JL, Koch SBJ, Veltman DJ, Olff M (2015): Reward functioning in PTSD: A systematic review exploring the mechanisms underlying anhedonia. Neurosci Biobehav Rev 51:189–204.
7. Husain M, Roiser JP (2018): Neuroscience of apathy and anhedonia: A transdiagnostic approach. Nat Rev Neurosci 19:470–484.
8. Pizzagalli DA (2022): Anhedonia: Preclinical, Translational, and Clinical Integration, vol. 58. Berlin: Springer Nature.
9. Guineau MG, Ikani N, Rinck M, Collard RM, van Eijndhoven P, Tendolkar I, *et al.* (2023): Anhedonia as a transdiagnostic symptom across psychological disorders: A network approach. Psychol Med 53:3908–3919.
10. Sutton RS, Barto AG (2018): Reinforcement Learning: An Introduction. Cambridge: MIT Press.
11. Chase HW, Frank MJ, Michael A, Bullmore ET, Sahakian BJ, Robbins TW (2010): Approach and avoidance learning in patients with major depression and healthy controls: Relation to anhedonia. Psychol Med 40:433–440.
12. Huys QJ, Pizzagalli DA, Bogdan R, Dayan P (2013): Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord 3:12.
13. Huys QJM, Browning M (2022): A Computational View on the Nature of Reward and Value in Anhedonia. Curr Top Behav Neurosci 58:421–441.
14. Parush N, Tishby N, Bergman H (2011): Dopaminergic balance between reward maximization and policy complexity. Front Syst Neurosci 5:22.
15. Gershman SJ (2020): Origin of perseveration in the trade-off between reward and complexity. Cognition 204:104394.
16. Lai L, Gershman SJ (2021): Policy compression: An information bottleneck in action selection. Psychol Learn Motivation 74:195–232.
17. Mikhael JG, Lai L, Gershman SJ (2021): Rational inattention and tonic dopamine. PLoS Comput Biol 17:e1008659.
18. Bari BA, Gershman SJ (2023): Undermatching is a consequence of policy compression. J Neurosci 43:447–457.
19. Manohar SG, Chong TTJ, Apps MAJ, Batla A, Stamelou M, Jarman PR, *et al.* (2015): Reward pays the cost of noise reduction in motor and cognitive control. Curr Biol 25:1707–1716.

20. Carlezon WA, Béguin C, DiNieri JA, Baumann MH, Richards MR, Todtenkopf MS, *et al.* (2006): Depressive-like effects of the kappa-opioid receptor agonist salvinorin A on behavior and neurochemistry in rats. J Pharmacol Exp Ther 316:440–447.
21. Bruijnzeel AW (2009): Kappa-opioid receptor signaling and brain reward function. Brain Res Rev 62:127–146.
22. Ebner SR, Roitman MF, Potter DN, Rachlin AB, Chartoff EH (2010): Depressive-like effects of the kappa Opioid receptor agonist salvinorin A are associated with decreased phasic dopamine release in the nucleus accumbens. Psychopharmacology 210:241–252.
23. Muschamp JW, Van't Veer A, Parsegian A, Gallo MS, Chen M, Neve RL, *et al.* (2011): Activation of creb in the nucleus accumbens shell produces anhedonia and resistance to extinction of fear in rats. J Neurosci 31:3095–3103.
24. Wallace CW, Holleran KM, Slinkard CY, Centanni SW, Jones SR (2024): Kappa opioid receptors negatively regulate real time spontaneous dopamine signals by reducing release and increasing uptake. bioRxiv https://doi.org/10.1101/2024.02.05.578840.
25. Tripp G, Alsop B (1999): Sensitivity to reward frequency in boys with attention deficit hyperactivity disorder. J Clin Child Psychol 28:366–375.
26. Pizzagalli DA, Jahn AL, O'Shea JP (2005): Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. Biol Psychiatry 57:319–327.
27. Krystal AD, Pizzagalli DA, Smoski M, Mathew SJ, Nurnberger J, Lisanby SH, *et al.* (2020): A randomized proof-of-mechanism trial applying the 'fast-fail' approach to evaluating kappa-opioid antagonism as a treatment for anhedonia. Nat Med 26:760–768.
28. Pizzagalli DA, Smoski M, Ang YS, Whitton AE, Sanacora G, Mathew SJ, *et al.* (2020): Selective kappa-opioid antagonism ameliorates anhedonic behavior: Evidence from the fast-fail trial in mood and anxiety spectrum disorders (fast-mas). Neuropsychopharmacology 45:1656–1663.
29. Gershman SJ, Lai L (2021): The reward-complexity trade-off in schizophrenia. Comput Psychiatr 5:38–53.
30. Lai L, Gershman SJ (2024): Human decision making balances reward maximization and policy compression. PLoS Comput Biol 20:e1012057.
31. Watkins CJ, Dayan P (1992): Q-learning. Mach Learn 8:279–292.
32. de Leeuw JR (2021): prt-test-my-brain. Available at: https://github.com/jodeleeuw/prt-test-my-brain. Accessed December 20, 2023.
33. Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M (2008): Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. J Psychiatr Res 43:76–87.
34. Pizzagalli DA, Goetz E, Ostacher M, Iosifescu DV, Perlis RH (2008): Euthymic patients with bipolar disorder show decreased reward learning in a probabilistic reward task. Biol Psychiatry 64:162–168.
35. Wise RA (2008): Dopamine and reward: The anhedonia hypothesis 30 years on. Neurotox Res 14:169–183.
36. Argyropoulos SV, Nutt DJ (2013): Anhedonia revisited: Is there a role for dopamine-targeting drugs for depression? J Psychopharmacol 27:869–877.
37. Pizzagalli DA, Evins AE, Schetter EC, Frank MJ, Pajtas PE, Santesso DL, Culhane M (2008): Single dose of a dopamine agonist impairs reinforcement learning in humans: Behavioral evidence from a laboratory-based measure of reward responsiveness. Psychopharmacology 196:221–232.
38. Berridge KC, Robinson TE (1998): What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28:309–369.
39. Rizvi SJ, Pizzagalli DA, Sproule BA, Kennedy SH (2016): Assessing anhedonia in depression: Potentials and pitfalls. Neurosci Biobehav Rev 65:21–35.
40. Strekalova T (2023): How the sucrose preference succeeds or fails as a measurement of anhedonia. In: Harro J, editor. Psychiatric Vulnerability, Mood, and Anxiety Disorders: Tests and Models in Mice and Rats. New York: Springer US, 89–102.
41. Wulff AB, Cooper P, Kodjo E, Abel E, Thompson SM (2023): How sucrose preference is gained and lost: An in-depth analysis of drinking behavior during the sucrose preference test in mice. eNeuro 10: ENEURO.0195-23.2023.
42. Rizvi SJ, Quilty LC, Sproule BA, Cyriac A, Michael Bagby R, Kennedy SH (2015): Development and validation of the dimensional anhedonia rating scale (DARS) in a community sample and individuals with major depression. Psychiatry Res 229:109–119.
43. Gard DE, Gard MG, Kring AM, John OP (2006): Anticipatory and consummatory components of the experience of pleasure: A scale development study. J Res Pers 40:1086–1102.
44. Khazanov GK, Ruscio AM, Forbes CN (2020): The positive valence systems scale: Development and validation. Assessment 27:1045–1069.
45. Robbins TW, Costa RM (2017): Habits. Curr Biol 27:R1200–R1206.
46. Miller KJ, Shenhav A, Ludvig EA (2019): Habits without values. Psychol Rev 126:292–311.
47. Nebe S, Kretzschmar A, Brandt MC, Tobler PN (2024): Characterizing human habits in the lab. Collabra Psychol 10:92949.
48. Laughlin SB (2001): Energy as a constraint on the coding and processing of sensory information. Curr Opin Neurobiol 11:475–480.
49. Bari BA, Gershman SJ (2024): Resource-rational psychopathology. Behav Neurosci 138:221–234.