Short Report

Chronic social stress induces isolated deficits in reward anticipation on a neuroeconomic foraging task

Romain Durand-de Cuttoli[1], Freddyson J. Martínez-Rivera[1], Long Li[1], Angélica Minier-Toribio[1], Scott J. Russo[1], Eric J. Nestler[1,2], Brian M. Sweis[1,2†]

Measuring reward anticipation distinct from other aspects of reward value, including costs required to obtain a reward or the intrinsic hedonic value of consuming the reward itself, can be difficult to disentangle. Here, we show that mice trained on a self-paced neuroeconomic foraging task convey reward anticipation via differentially invigorated travel times between uniquely flavored feeding sites separate from willingness to wait, consummatory behaviors, or place preferences measured within the same trial. Following exposure to chronic social defeat stress, we found that only stress-susceptible but not stress-resilient mice revealed deficits in this metric after consuming but not after rejecting a reward on the previous trial, indicating that blunted anticipation in these animals is state-dependent, or punctuated by recent reward receipt. After increasing economic pressure and task demands, locomotion was globally invigorated and, in turn, masked stress-related deficits in reward anticipation. These findings suggest that the ability to detect changes in specific aspects of motivational deficits associated with depression and other stress-related disorders depends on an interaction between the state of an individual and environmental circumstances.

Anhedonia describes a loss of interest or pleasure in rewarding activities and can manifest in the form of decreased motivation to engage in reward-seeking behaviors (1). Deficits in reward anticipation comprise one aspect of anhedonia. Levels of motivation, however, are highly dynamic and can be altered, for example, due to fluctuations in different physiological states of an individual or as a result of environmental circumstances (2).

Animal studies operationalize anhedonia in various ways, including decreased sucrose preference, impaired conditioned place preference, diminished effort expenditures on operant-based tasks (e.g., lever pressing or nose poking for food), blunted sensitivity on probabilistic reward tasks, reduced willingness to wait on delayed discounting tasks, and even alterations in orofacial expressions following reward consumption (2-11). Such experimental approaches, while instrumental in laying a foundational understanding of the neurobiology of motivation, include aspects of reward anticipation that are inherently confounded by hedonic processes related to the appraisal of the consumption of the reward itself or are directly linked to variations in the cost required to obtain the reward. It is therefore difficult to disentangle multiple properties of reward value from isolated, independent measures of intrinsic motivation pertaining solely to the process of reward anticipation.

Here, we tested mice previously exposed to chronic social defeat stress, a well-validated animal model used for the study of depression (12-15), on a translational neuroeconomic self-paced foraging paradigm, termed "Restaurant Row" (12-21). This task provides a rich framework to operationalize reward value and motivation-related processes among several domains of complex behavior separated across space and time. By analyzing travel behavior between reward patches, we found that the speed with which mice traveled was invigorated when approaching more preferred rewards – a metric of reward anticipation that was independent of other aspects of reward value measured within the same trial. Reward anticipation was uniquely disrupted in stress-susceptible but not stress-resilient mice, compared to non-defeated controls, but such deficits depended upon whether or not another reward was recently obtained as well as the level of economic demand of the task. These findings suggest that the ability to detect motivational deficits after chronic stress is state-specific and can be masked by environmental pressure.

Results

Following exposure to chronic social defeat stress (10 consecutive days, Fig. 1a-b), C57BL/6J male mice were categorized into either stress-susceptible (SUS) or stress-resilient (RES) groups based on a brief social interaction assay that has served as a well-validated predictor of several depression-related phenotypes on other rapid screening tests and on some more extensive behavioral procedures (Fig. 1d) (22-29). Next, these animals alongside non-defeated controls (CON) were tested across 55 consecutive days on Restaurant Row longitudinally (Fig. 1a), during which animals spend a limited time budget foraging for their only source of food (Fig. 1c) (12-15). On this task, mice make serial decisions accepting or rejecting offers presented for food rewards. The maze is divided into four uniquely flavored and contextualized feeding sites, or "restaurant" patches. Flavors are used to manipulate reward value without assuming individual differences in subjective preferences and without introducing variable handling time (e.g., as opposed to introducing different reward sizes). Each restaurant contains a separate offer zone and wait zone. Upon entry into a restaurant's offer zone, a tone sounded in the offer

zone whose pitch indicated the delay mice would have to wait in a cued countdown in order to earn a pellet should mice choose to enter the wait zone. After consuming an earned reward in the wait zone or skipping an offer in the offer zone, mice were required to advance to the next restaurant.

On days 1-5 of the task, all offers were only 1 s in length. During this time, mice displayed stable subjective flavor preferences that were ranked by summing end of session earnings in each restaurant and were no different on average between CON, RES, and SUS mice (Fig. 1f), running a similar number of laps (Fig. 1e) and stabilizing at equivalent body weights (Fig. 1g). Next, mice advanced into an increasingly reward scarce environment every five days, being exposed to a more expensive range of offers (days 6-10: 1 to 5 s offers; days 11-15: 1 to 15 s offers, days 16+: 1 to 30 s offers, Fig. 1a). While the probability of entering the
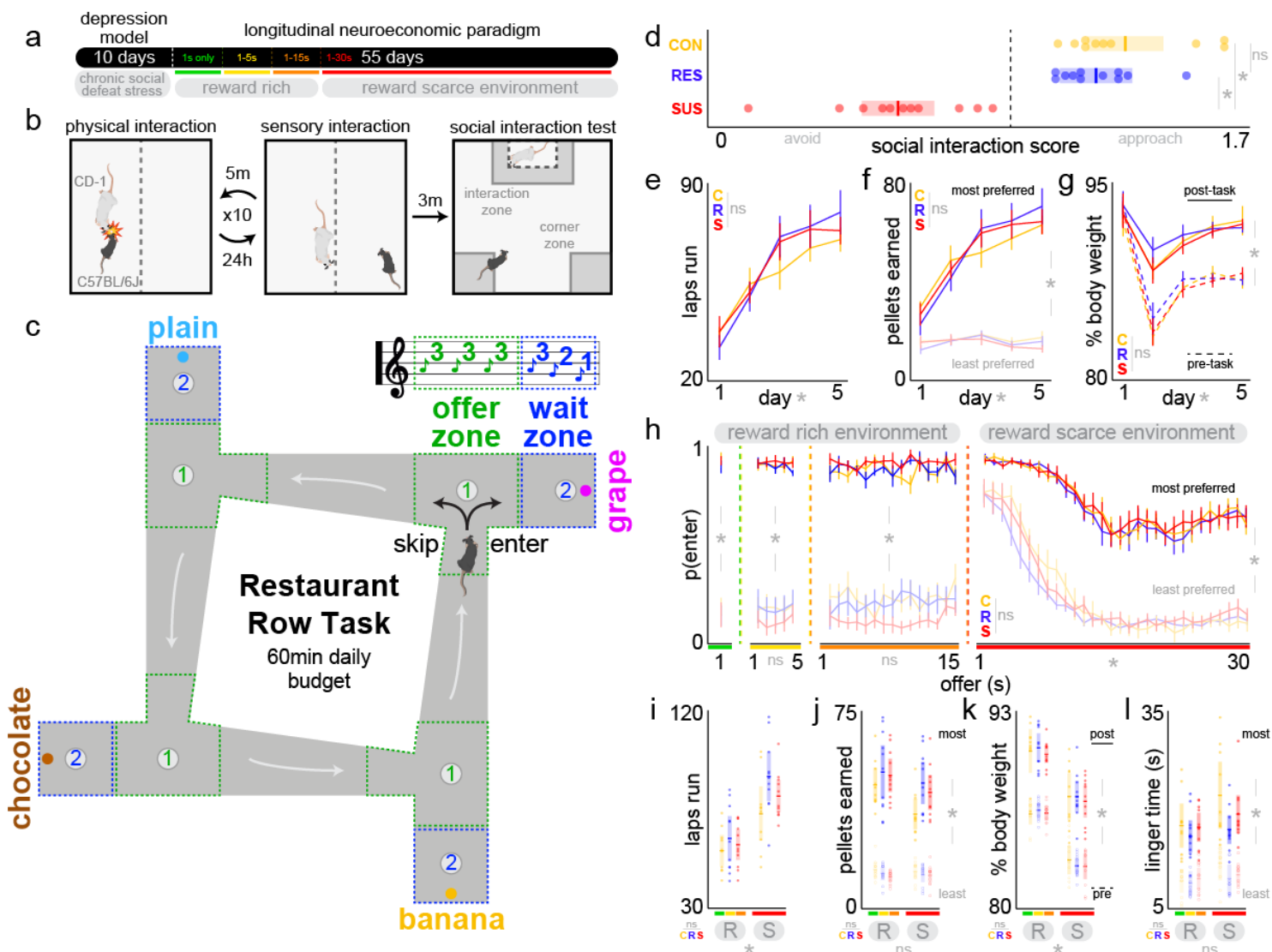

Figure 1. Mice exposed to chronic social defeat stress are capable of learning the Restaurant Row task without overt differences in revealed subjective flavor preferences, economic decisions, or post-consumption hedonic valuations. (a) Experimental timeline. (b) Chronic social defeat stress protocol. C57BL/6J mice are exposed to different aggressive CD-1 mice for 10 consecutive days (5-minute daily physical interactions followed by co-housing with a divider for the rest of the day). After the final defeat exposure, C57BL/6J mice are screened on the social interaction test. Time spent in the interaction zone is quantified before (3-minute baseline probe) and after (3-minute test probe) a novel CD-1 target mouse is placed in a small interaction chamber. (c) Restaurant Row task schematic. Mice have a limited time budget traversing a square maze with four uniquely flavored and contextualized restaurants, making serial accept versus reject decisions encountering offers randomly selected from a uniform distribution of delayed rewards cued by tone pitch in the offer zone that counts down should mice choose to enter the wait zone. (d) Social interaction score derived during the social interaction test is defined by ratio of time spent in the interaction zone when the CD-1 target is present versus when absent. Vertical dashed line represents score of 1.0. Data below this line reflect avoidance behavior that was used to define stress-susceptible (SUS, n=10) mice versus stress-resilient (RES, n=11) mice who instead approach the novel CD-1 target mouse during the social interaction screen, like non-defeated controls (CON, n=11). One-way ANOVA: $F_{2,31}$=47.975, *p<0.0001, Tukey$_{CON/RES}$ t=1.15, ^{ns}p=0.495. (e) Laps run in the correct direction. Two-way ANOVA: Significant main effect across days: $F_{4,31}$=81.953, *p<0.0001; no significant interaction with group: $F_{8,159}$=1.676, ^{ns}p=0.19. (f) Pellets earned in most preferred and least preferred restaurants. Rankings defined by summing end of session earns. Three-way ANOVA: Significant interaction between day and rank: $F_{4,139}$=78.771, *p<0.0001; no significant interaction between day, rank, and group: $F_{4,319}$=1.207, ^{ns}p=0.30. (g) Percent bodyweight measured before and after testing relative to baseline weight defined before day 1 of Restaurant Row testing. Three-way ANOVA: Significant interaction between day and pre- vs. post-task measurements: F $F_{4,319}$=9.966, *p<0.01; no significant interaction between day, pre- vs. post-task measurements, and group: $F_{4,319}$=0.017, ^{ns}p=0.98. (h) Probability of accepting an offer as a function of cued offer cost in a reward rich environment (1 s only [green], 1 to 5 s [yellow], and 1 to 15 s [orange] offers, note mice accept every offer in this environment in the most preferred restaurant regardless of cost) versus a reward scarce environment (1 to 30 s [red] offers). Four-way ANOVA: Significant interaction between offer, rank, and block of training: $F_{87,7038}$=6.476, *p<0.01; no significant interaction between offer, rank, block of training, and group: $F_{174,7038}$=0.001, ^{ns}p=0.99. (i-l) Behaviors compared between reward rich (R) and reward scarce (S) environments. (i) laps run in the correct direction. Two-way ANOVA: Significant main effect across environments: $F_{1,31}$=42.850, *p<0.0001; no significant interaction with group: $F_{2,63}$=0.910, ^{ns}p=0.41. (j) Pellets earned in most and least preferred restaurants. Three-way ANOVA: No significant main effect across environments: $F_{1,31}$=0.012, ^{ns}p=0.91; significant main effect of rank: $F_{1,31}$=524.843, *p<0.0001; no significant interaction between environment, rank, and group: $F_{2,127}$=0.456, ^{ns}p=0.63. (k) Percent baseline bodyweight measured before and after testing. Three-way ANOVA: Significant main effect across environments: $F_{1,31}$=188.431, *p<0.0001; significant main effect of pre- vs. post-task measurements: $F_{1,31}$=312.310, *p<0.0001; no significant interaction between environment, pre- vs. post-task, and group: $F_{2,127}$=0.077, ^{ns}p=0.93. (l) Time spent lingering at the reward site after consuming an earned reward before advancing to the next trial in most and least preferred restaurants. Three-way ANOVA: No significant main effect across environments: $F_{1,31}$=2.034, ^{ns}p=0.50; significant main effect of rank: $F_{1,31}$=83.936, *p<0.0001; no significant interaction between environments, rank, and group: $F_{2,127}$=0.118, ^{ns}p=0.89. Error bars represent ±1 SEM. Dots represent individual animals.

Durand-de Cuttoli et al. (2022). Chronic social stress induces isolated deficits in reward anticipation on a neuroeconomic foraging task.
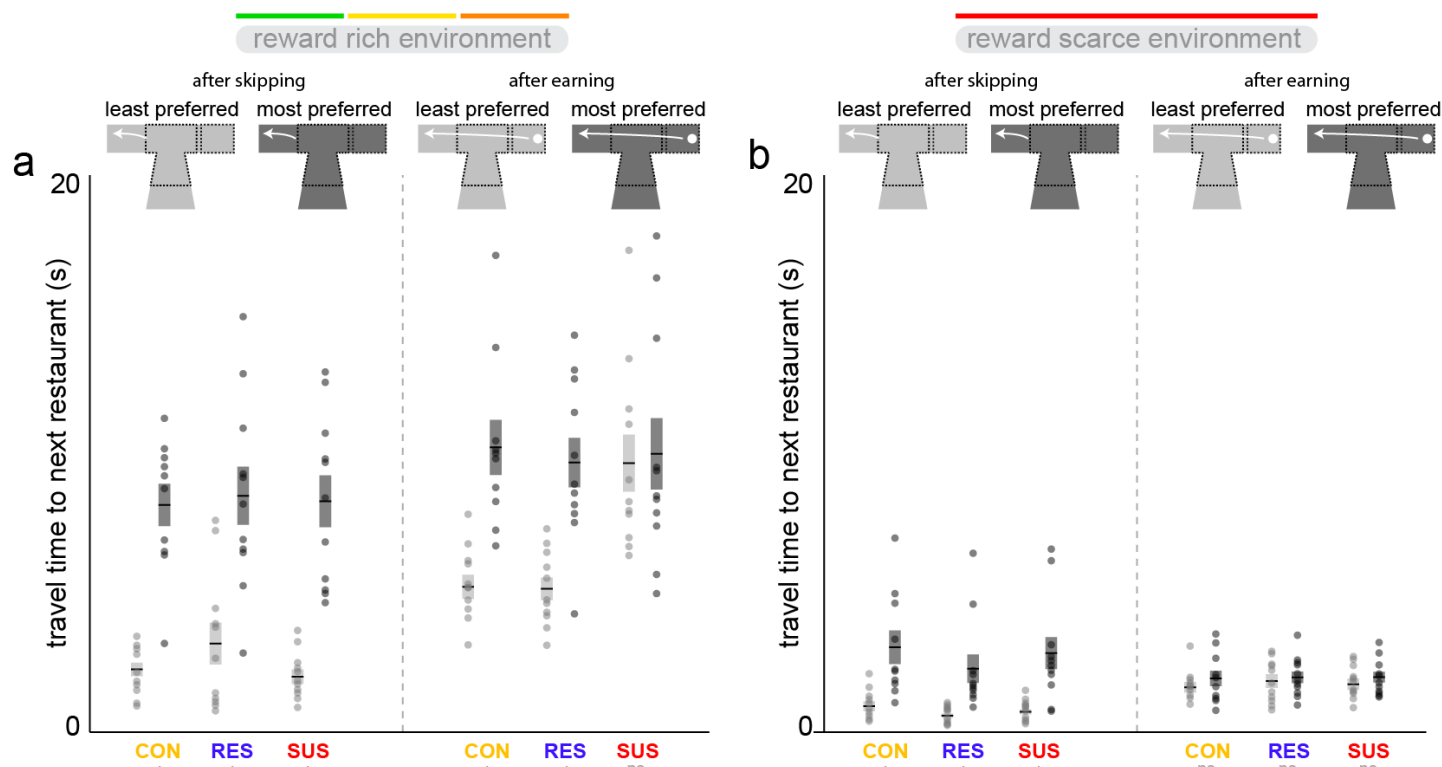
Figure 2. Travel time between restaurants reflects anticipatory behavior with a deficit in SUS mice that is state-specific and sensitive to economic pressure. (a) Travel time between restaurants in a reward rich environment (1 s only [green], 1 to 5 s [yellow], and 1 to 15 s [orange] offers) after a skip decision was made in the offer zone (left) versus after an earned reward was consumed and mice departed the wait zone (right) separated by the ranking of which restaurant is being departed. Overall, mice travel faster to the next restaurant after departing the least preferred restaurant (light gray) compared to after departing the most preferred restaurant (dark gray), main effect of rank: $F_{1,31}$=90.264, *p<0.0001. Despite all animals showing this invigorated behavioral response after skipping a previous offer in the least preferred restaurant compared to the most preferred restaurant, only SUS mice do not display this invigorated behavioral response after earning rewards (significant interaction between rank, enter vs. skip decision on the previous trial, and group: $F_{2,127}$=3.318, *p<0.05). (b) Travel time between restaurants in a reward scarce environment (1 to 30 s [red] offers). Overall mice travel faster in a reward scarce environment compared to a reward rich environment: $F_{1,31}$=87.189, *p<0.0001. In this environment, invigorated travel times were observable after skipping the least preferred restaurant compared to after skipping the most preferred restaurant (left, $F_{1,31}$=35.811, *p<0.0001), with no interaction with group ($F_{2,63}$=0.202, ^{ns}p=0.82). However, in this environment, there were no detectable differences in travel times observed after earning in either the least or most preferred restaurants in all groups (right, no main effect of rank: $F_{1,31}$=2.060, ^{ns}p=0.16, no interaction with group: $F_{2,63}$=0.135, ^{ns}p=0.87).

wait zone was consistently higher for more preferred restaurants compared to less preferred restaurants across the entire experiment, mice did not factor in the cued cost of the offer in their offer zone decisions until entering the 1 to 30 s block of testing (Fig. 1h). CON, RES, and SUS mice did not differ in their ability to effectively discriminate tones in the offer zone during this block. By adopting this change in strategy in a reward scarce environment, mice were able to maintain food intake (Fig. 1j) but in turn expended more energy by running more laps (Fig. 1i) and thus had lower body weights (Fig. 1k), with no differences between CON, RES, and SUS mice. Furthermore, by measuring the time after consuming earned rewards until departing the wait zone, we found that mice lingered at the reward site before advancing to the next trial, lingering longer in more preferred restaurants (Fig. 1l). This metric represents a post-consumption within-trial place-preference associated with the restaurant's context that was stable across the entire experiment and was not different among CON, RES, and SUS mice.

While there were no overt differences among CON, RES, and SUS mice in several measures of reward value, how mice approached each restaurant revealed other aspects of motivation that were different between groups. Inter-trial travel time was measured in the corridor between restaurants either from when a skip decision was made in the offer zone or from when mice departed the wait zone after consuming an earned reward. By analyzing travel time following skip decisions in a reward rich environment, we found that all mice arrived at the next restaurant more quickly when departing the least preferred restaurant compared to the most preferred restaurant (Fig. 2a). These data suggest that travel time captures a form of reward anticipation and invigorated motivational state in between trials based solely on the identity of the

upcoming restaurant signaled by spatial context and independent of the delay of the next reward offer, which is not cued until arriving at the next restaurant.

Next, by analyzing travel time after earns, overall travel times were slower compared to after skipping. Nonetheless, anticipatory effects were still measurable, with faster travel times once departing after earning in least preferred restaurants compared to most preferred restaurants in CON and RES mice (Fig. 2a). Interestingly, this effect was absent in SUS mice such that travel times were just as slow when departing least preferred restaurants after consuming an earned reward compared to when departing most preferred restaurants. These data indicate that anticipatory responses of more preferred future reward opportunities in SUS mice is blunted, not globally because anticipation is intact when skipping, but only upon reward receipt in least preferred restaurants. This suggests that impairments in eagerness to travel in SUS mice is state-specific. "State" here is defined by whether or not food was consumed on the previous trial. This finding highlights how the receipt of a reward only in SUS mice serves as an experience capable of devaluing the anticipation of better, future rewards while other animals under similar circumstances are able to invigorate their behavioral response in pursuit of upcoming opportunities.

Compared to behavior in a reward rich environment, travel time in a reward scarce environment overall was significantly faster (Fig. 2b) as mice were pressured to run more laps and make faster, more economically advantageous choices similar for CON, RES, and SUS mice. Despite this overall increase in speed, anticipation was still observable when skipping least preferred restaurants compared to most preferred restaurants in CON, RES, and SUS mice. Interestingly, anticipation after

consuming an earned reward was no longer detectable in any group. These data suggest that the level of external economic pressure placed on animals can increase demand on foraging locomotion. This pressure can in turn mask state-specific anticipation and conceal potential stress-induced deficits in motivation.

Discussion

Anticipating a reward can alter behavior in numerous ways, from increasing salivation in classical Pavlovian conditioning to invigorating more complex goal-oriented actions through instrumental transfer (*30, 31*). Deficits in reward anticipation can lead to impairments in motivation often observed in stress-related disorders such as depression, but it is unclear if such changes are linked to the intrinsic value of the reward itself, work required to obtain the reward, or other processes reflecting interactions between various internal states and external circumstances (*32*). We measured reward anticipation by examining travel behavior between serial foraging decisions, which is often formalized into neuroeconomic models of energetic optimization but is typically held constant, controlled for, or ignored in lab experiments rather than viewed as a dependent measure of interest (*33-37*). Here, we report stress-induced deficits in eagerness to travel that can be isolated in space and time from other decision- or consummatory-related valuations measured within the same trial and are dependent on reward history and environmental pressure.

Invigorated locomotion has previously been used to capture motivation but is often measured in tasks that do not parametrically vary value along multiple dimensions or different action-selection processes (*2, 38*). Furthermore, such tasks typically measure the pursuit of a single reward in an operant box with anticipation behavior operationalized by the degree of lever pressing, nose poking, or lick rate at a reward port (*39-43*). These metrics are inherently linked to preparatory consummatory behavior, goal tracking, or temporal error of estimated reward delivery (*44, 45*). Nonetheless, investigating the neural correlates of such processes have revealed a critical role of dopaminergic transmission interacting with reward circuitry and psychomotor systems. Dopamine signaling has been shown to respond to a variety of events on a spectrum from pure reward to pure movement, including unconditioned reward consumption, cue-related reward prediction, movement-locked reward-seeking behavior, generalized reward-related invigoration, and strict locomotor movements, broadly tuned to a variety of actions and typically bursting just before acceleration (*46-54*). Dopamine signaling is thought to gate the excitability of striatal output

neurons that integrate cortical and thalamic reward-related action plans from coincident glutamatergic inputs translated into motor events (*55, 56*). Although much of the literature on dysfunction in such processes has traditionally focused on pathology involved in Parkinson's disease, schizophrenia, and addiction, several studies have explored how alterations in dopamine transmission may underlie motivational deficits in stress-related disorders (*57*). Blunted reward prediction error signals in the striatum and ventral tegmental area have been observed in patients with major depressive disorder as well as in animals following exposure to stress (*26, 58-61*). Acute stress has been shown to alter the neural encoding of reward anticipation associated with diminished anticipatory licking behavior (*39*). Chronic social defeat stress alters dopaminergic firing rates in SUS mice, linked to anhedonia-related phenotypes but not reward anticipation processes specifically (*26, 60*). In the present study, why anticipatory behavior in SUS mice remains intact when skipping but not after recently earning a reward is perplexing. We suggest that state-specific differences here likely involve other neurotransmitter systems such as serotonin, which can signal quiescence of reward-related vigor following food consumption and have more classically been implicated in depression (*62-67*). Consuming a reward may involve a "reset" in relative value state from which a stress-related deficit emerges that is not present in the computations during skipping (Fig. 3). Dopamine and serotonin transmission is likely interacting here, both of which have been shown to contribute to more complex foraging strategies that incorporate travel speed, hunger, energy optimization, and opportunity costs when exploring an environment (*68, 69*). Given enough task pressure, decision history may have negligible effects on the physiological processes underlying movement vigor and may be less sensitive to reveal stress-related dysfunction.

We isolated a measure of reward anticipation based on travel vigor during a neuroeconomic task. We discovered that stress-susceptible but not stress-resilient mice show impairments in this measure of reward anticipation not merely as a result of a global deficit in anticipation capabilities but rather only following reward consumption. These data provide a novel perspective to interrogate how a rewarding stimulus only in a depressive state could devalue the anticipation of future rewards (*70, 71*). We also highlight the importance of careful considerations of task demands in being able to extract stress-related differences in motivation. Our approach can shed light on more nuanced aspects of anhedonia, unpacking how state and environment interact in certain circumstances to influence the way depressed individuals struggle with motivational inertia and perceive what may lie on the horizon ahead.

Figure 3. Summary of depression-related deficits in reward anticipation that are state-specific and masked by economic demand. Travel time between restaurants when traveling from the least preferred restaurant to more preferred restaurants, defined by rankings of total earns in each restaurant, and depicted by the transition in contrast from light to dark gray, is decreased (or invigorated, represented by the chevrons) relative to departures in the reverse value sequence (when traveling from the most preferred restaurant to less preferred restaurants, represented by the thin gradient bar from dark to light gray located beneath the corridor). This anticipatory effect is present in non-defeated control mice (C), stress-resilient mice (R), and stress-susceptible mice (S) after skipping an offer presented in the least preferred restaurant while in an economically rich environment (low economic pressure, top left). This anticipatory effect is also present in C and R mice after earning a reward in this environment (chevrons present), but is absent in S mice (chevrons absent, top right). In an economically scarce environment (high economic pressure, bottom), all mice travel faster (represented by the increase in chevrons). Anticipatory effects are still present in C, R, and S mice after skipping the least preferred restaurant (bottom left) but are no longer detectable in any group (represented by the dashed lines, increased spacing reflects the overall faster travel speed in this environment, bottom right).
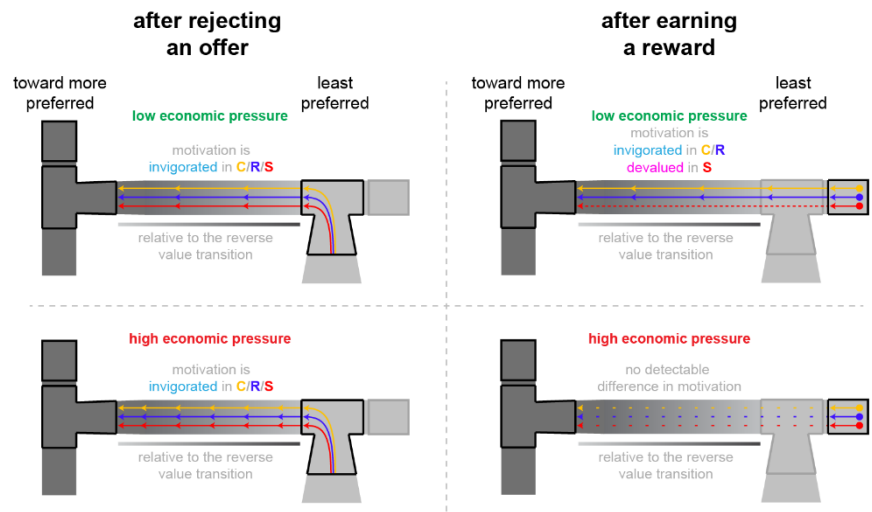
References

1. American Psychiatric Association., American Psychiatric Association. DSM-5 Task Force., *Diagnostic and statistical manual of mental disorders : DSM-5.* (American Psychiatric Association, Washington, D.C., ed. 5th, 2013), pp. xliv, 947 p.
2. K. C. Berridge, Motivation concepts in behavioral neuroscience. *Physiol Behav* **81**, 179-209 (2004).
3. K. C. Berridge, T. E. Robinson, J. W. Aldridge, Dissecting components of reward: 'liking', 'wanting', and learning. *Curr Opin Pharmacol* **9**, 65-73 (2009).
4. A. Der-Avakian, S. A. Barnes, A. Markou, D. A. Pizzagalli, Translational Assessment of Reward and Motivational Deficits in Psychiatric Disorders. *Curr Top Behav Neurosci* **28**, 231-262 (2016).
5. A. Der-Avakian, M. S. D'Souza, D. A. Pizzagalli, A. Markou, Assessment of reward responsiveness in the response bias probabilistic reward task in rats: implications for cross-species translational research. *Transl Psychiatry* **3**, e297 (2013).
6. E. Acquas, E. Carboni, P. Leone, G. Di Chiara, SCH 23390 blocks drug-conditioned place-preference and place-aversion: anhedonia (lack of reward) or apathy (lack of motivation) after dopamine-receptor blockade? *Psychopharmacology (Berl)* **99**, 151-155 (1989).
7. M. Papp, P. Willner, R. Muscat, An animal model of anhedonia: attenuation of sucrose consumption and place preference conditioning by chronic unpredictable mild stress. *Psychopharmacology (Berl)* **104**, 255-259 (1991).
8. B. D. Kangas *et al.*, A cross-species assay demonstrates that reward responsiveness is enduringly impacted by adverse, unpredictable early-life experiences. *Neuropsychopharmacology*, (2021).
9. N. Shafiei, M. Gray, V. Viau, S. B. Floresco, Acute stress induces selective alterations in cost/benefit decision-making. *Neuropsychopharmacology* **37**, 2194-2209 (2012).
10. E. Ordones Sanchez *et al.*, Early life adversity promotes resilience to opioid addiction-related phenotypes in male rats and sex-specific transcriptional changes. *Proc Natl Acad Sci U S A* **118**, (2021).
11. V. Vialou *et al.*, Prefrontal cortical circuit for depression- and anxiety-related behaviors mediated by cholecystokinin: role of DeltaFosB. *J Neurosci* **34**, 3878-3887 (2014).
12. B. M. Sweis *et al.*, Sensitivity to "sunk costs" in mice, rats, and humans. *Science* **361**, 178-181 (2018).
13. B. M. Sweis, E. B. Larson, A. D. Redish, M. J. Thomas, Altering gain of the infralimbic-to-accumbens shell circuit alters economically dissociable decision-making algorithms. *Proc Natl Acad Sci U S A* **115**, E6347-E6355 (2018).
14. B. M. Sweis, A. D. Redish, M. J. Thomas, Prolonged abstinence from cocaine or morphine disrupts separable valuations during decision conflict. *Nat Commun* **9**, 2521 (2018).
15. B. M. Sweis, M. J. Thomas, A. D. Redish, Mice learn to avoid regret. *PLoS Biol* **16**, e2005853 (2018).
16. A. P. Steiner, A. D. Redish, Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat Neurosci* **17**, 995-1002 (2014).
17. S. V. Abram, Y. A. Breton, B. Schmidt, A. D. Redish, A. W. MacDonald, 3rd, The Web-Surf Task: A translational model of human decision-making. *Cogn Affect Behav Neurosci* **16**, 37-50 (2016).
18. S. V. Abram, M. Hanke, A. D. Redish, A. W. MacDonald, 3rd, Neural signatures underlying deliberation in human foraging decisions. *Cogn Affect Behav Neurosci* **19**, 1492-1508 (2019).
19. S. V. Abram, A. D. Redish, A. W. MacDonald, 3rd, Learning From Loss After Risk: Dissociating Reward Pursuit and Reward Valuation in a Naturalistic Foraging Task. *Front Psychiatry* **10**, 359 (2019).
20. T. Huynh, K. Alstatt, S. V. Abram, N. Schmitzer-Torbert, Vicarious Trial-and-Error Is Enhanced During Deliberation in Human Virtual Navigation in a Translational Foraging Task. *Front Behav Neurosci* **15**, 586159 (2021).
21. R. Kazinka, A. W. MacDonald, 3rd, A. D. Redish, Sensitivity to Sunk Costs Depends on Attention to the Delay. *Front Psychol* **12**, 604843 (2021).
22. A. Der-Avakian *et al.*, Social defeat disrupts reward learning and potentiates striatal nociceptin/orphanin FQ mRNA in rats. *Psychopharmacology (Berl)* **234**, 1603-1614 (2017).
23. A. Der-Avakian, M. S. Mazei-Robison, J. P. Kesby, E. J. Nestler, A. Markou, Enduring deficits in brain reward function after chronic social defeat in rats: susceptibility, resilience, and antidepressant response. *Biol Psychiatry* **76**, 542-549 (2014).
24. H. Kronman *et al.*, Long-term behavioral and cell-type-specific molecular effects of early life stress are mediated by H3K79me2 dynamics in medium spiny neurons. *Nat Neurosci* **24**, 667-676 (2021).
25. Z. S. Lorsch *et al.*, Stress resilience is promoted by a Zfp189-driven transcriptional network in prefrontal cortex. *Nat Neurosci* **22**, 1413-1423 (2019).
26. V. Krishnan *et al.*, Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391-404 (2007).
27. H. Choudhury *et al.*, Improvement of cellular uptake, in vitro antitumor activity and sustained release profile with increased bioavailability from a nanoemulsion platform. *Int J Pharm* **460**, 131-143 (2014).
28. R. C. Bagot *et al.*, Circuit-wide Transcriptional Profiling Reveals Brain Region-Specific Gene Networks Regulating Depression Susceptibility. *Neuron* **90**, 969-983 (2016).
29. C. J. Pena *et al.*, Early life stress confers lifelong stress susceptibility in mice via ventral tegmental area OTX2. *Science* **356**, 1185-1188 (2017).
30. M. S. Fanselow, K. M. Wassum, The Origins and Organization of Vertebrate Pavlovian Conditioning. *Cold Spring Harb Perspect Biol* **8**, a021717 (2015).
31. E. Cartoni, B. Balleine, G. Baldassarre, Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci Biobehav Rev* **71**, 829-848 (2016).
32. D. A. Pizzagalli, Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annu Rev Clin Psychol* **10**, 393-423 (2014).
33. T. Kalenscher, C. M. Pennartz, Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Prog Neurobiol* **84**, 284-315 (2008).
34. E. L. Charnov, Optimal foraging, the marginal value theorem. *Theor Popul Biol* **9**, 129-136 (1976).
35. A. M. Wikenheiser, D. W. Stephens, A. D. Redish, Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proc Natl Acad Sci U S A* **110**, 8308-8313 (2013).
36. E. C. Carter, A. D. Redish, Rats value time differently on equivalent foraging and delay-discounting tasks. *J Exp Psychol Gen* **145**, 1093-1101 (2016).
37. D. W. Stephens, Decision ecology: foraging and the ecology of animal decision making. *Cogn Affect Behav Neurosci* **8**, 475-484 (2008).
38. Y. A. Breton, A. Mullett, K. Conover, P. Shizgal, Validation and extension of the reward-mountain model. *Front Behav Neurosci* **7**, 125 (2013).
39. D. C. Lowes *et al.*, Ventral tegmental area GABA neurons mediate stress-induced blunted reward-seeking in mice. *Nat Commun* **12**, 3539 (2021).
40. K. I. Bakhurin *et al.*, Differential Encoding of Time by Prefrontal and Striatal Network Dynamics. *J Neurosci* **37**, 854-870 (2017).
41. T. W. Chen, N. Li, K. Daie, K. Svoboda, A Map of Anticipatory Activity in Mouse Motor Cortex. *Neuron* **94**, 866-879 e864 (2017).
42. T. Spellman, M. Svei, J. Kaminsky, G. Manzano-Nieves, C. Liston, Prefrontal deep projection neurons enable cognitive flexibility via persistent feedback monitoring. *Cell* **184**, 2750-2766 e2717 (2021).
43. C. M. Constantinople *et al.*, Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *Elife* **8**, (2019).
44. S. B. Flagel, S. J. Watson, T. E. Robinson, H. Akil, Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology (Berl)* **191**, 599-607 (2007).
45. C. K. Starkweather, B. M. Babayan, N. Uchida, S. J. Gershman, Dopamine reward prediction errors reflect hidden-state inference across time. *Nat Neurosci* **20**, 581-589 (2017).
46. B. T. Saunders, J. M. Richard, E. B. Margolis, P. H. Janak, Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat Neurosci* **21**, 1072-1083 (2018).
47. T. N. Lerner, A. L. Holloway, J. L. Seiler, Dopamine, Updated: Reward Prediction Error and Beyond. *Curr Opin Neurobiol* **67**, 123-130 (2021).
48. J. D. Berke, What does dopamine mean? *Nat Neurosci* **21**, 787-793 (2018).
49. W. Schultz, Predictive reward signal of dopamine neurons. *J Neurophysiol* **80**, 1-27 (1998).
50. A. V. Kravitz, A. C. Kreitzer, Striatal mechanisms underlying movement, reinforcement, and punishment. *Physiology (Bethesda)* **27**, 167-177 (2012).
51. A. V. Kravitz *et al.*, Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622-626 (2010).
52. A. L. Wong, M. A. Lindquist, A. M. Haith, J. W. Krakauer, Explicit knowledge enhances motor vigor and performance: motivation versus practice in sequence tasks. *J Neurophysiol* **114**, 219-232 (2015).

53. C. Spielewoy *et al.*, Behavioural disturbances associated with hyperdopaminergia in dopamine-transporter knockout mice. *Behav Pharmacol* **11**, 279-290 (2000).
54. K. M. Wassum, S. B. Ostlund, N. T. Maidment, Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol Psychiatry* **71**, 846-854 (2012).
55. F. Tecuapetla, X. Jin, S. Q. Lima, R. M. Costa, Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. *Cell* **166**, 703-715 (2016).
56. D. Thura, P. Cisek, The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. *Neuron* **95**, 1160-1170 e1165 (2017).
57. P. Mazzoni, A. Hristova, J. W. Krakauer, Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* **27**, 7105-7116 (2007).
58. P. Kumar *et al.*, Impaired reward prediction error encoding and striatal-midbrain connectivity in depression. *Neuropsychopharmacology* **43**, 1581-1588 (2018).
59. K. M. Tye *et al.*, Dopamine neurons modulate neural encoding and expression of depression-related behaviour. *Nature* **493**, 537-541 (2013).
60. D. Chaudhury *et al.*, Rapid regulation of depression-related behaviours by control of midbrain dopamine neurons. *Nature* **493**, 532-536 (2013).
61. J. Barik *et al.*, Chronic stress triggers social aversion via glucocorticoid receptor in dopaminoceptive neurons. *Science* **339**, 332-335 (2013).
62. Y. Li *et al.*, Synaptic mechanism underlying serotonin modulation of transition to cocaine addiction. *Science* **373**, 1252-1256 (2021).
63. T. R. Gruninger, B. LeBoeuf, Y. Liu, L. R. Garcia, Molecular signaling involved in regulating feeding and other motivated behaviors. *Mol Neurobiol* **35**, 1-20 (2007).
64. R. Cools, K. Nakamura, N. D. Daw, Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology* **36**, 98-113 (2011).
65. Y. L. Boureau, P. Dayan, Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology* **36**, 74-97 (2011).
66. R. Cools, A. C. Roberts, T. W. Robbins, Serotoninergic regulation of emotional and behavioural control processes. *Trends Cogn Sci* **12**, 31-40 (2008).
67. C. Varazzani, A. San-Galli, S. Gilardeau, S. Bouret, Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J Neurosci* **35**, 7866-7877 (2015).
68. Y. Niv, N. D. Daw, D. Joel, P. Dayan, Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* **191**, 507-520 (2007).
69. E. Lottem *et al.*, Activation of serotonin neurons promotes active persistence in a probabilistic foraging task. *Nat Commun* **9**, 1000 (2018).
70. B. Balleine, A. Dickinson, Signalling and incentive processes in instrumental reinforcer devaluation. *Q J Exp Psychol B* **45**, 285-301 (1992).
71. R. M. Carelli, E. A. West, When a good taste turns bad: Neural mechanisms underlying the emergence of negative affect and associated natural reward devaluation by cocaine. *Neuropharmacology* **76**, 360-369 (2014).

Author contributions

Conceptualization: BMS; Methodology: RDC, SJR, EJN, BMS; Investigation: RDC, FMR, LL, AMT, BMS; Data curation: RDC, BMS; Formal analysis: RDC, BMS; Funding acquisition: SJR, EJN; Supervision: RDC, SJR, EJN, BMS; Writing – original draft: RDC, BMS; Writing – review & editing: all authors

Data and materials availability

All data, code, and materials used in the analysis are available in the materials and methods section or upon request.

Competing interests

Authors declare that they have no competing interests.

Materials and methods

Animals and husbandry: We purchased 10-week-old wild-type male C57BL/6J mice (Jackson Laboratory). Additionally, we purchased 16 to 24-week-old male sexually experienced retired breeder CD-1 (ICR) mice (Charles River Laboratories) that were used as aggressors for the chronic social defeat stress protocol. All mice were maintained on a 12-hr light/dark cycle with ad libitum access to water. Experiments were conducted during the light phase. Experiments were approved by the Mount Sinai Institutional Animal Care and Use Committee (IACUC; protocol number LA12-00051) and adhered to the National Institutes of Health (NIH) guidelines.

Chronic Social Defeat Stress: A single C57BL/6J mouse was co-housed with a single CD-1 mouse and allowed to experience aggression behavior for 5-10 min of attacking before being separated by a mesh divider for the remainder of the day. Mice then no longer had direct physical contact but continued to have visual, olfactory, and auditory contact. This was repeated for 10 consecutive days with 10 different CD-1 mice. Non-defeated control C57BL/6J mice were exposed instead to other, domiciled C57BL/6J mice. Mice had access to regular chow ad libitum during this protocol and were weighed daily.

Social Interaction Screen: After the defeat protocol, mice were individually housed and assayed on a social interaction test that can capture social avoidance. A single C57BL/6J mouse was placed in a large open field arena with a novel CD-1 mouse enclosed in a small chamber. EthoVision software was used to track the location of the C57BL/6J mouse during this social interaction assay. Time spent near (interaction zone) versus away from the CD-1 mouse was used to quantify a social interaction score calculated from time in the interaction zone with the CD-1 mouse present in the chamber (2.5-min trial) relative to a preceding 2.5-min baseline trial without the target CD-1 mouse present.

Neuroeconomic Decision-Making Paradigm: Restaurant Row Task: After the social interaction screen, mice were switched to a full-nutrition flavored pellet diet (BioServe products; 20 mg dustless precision full-nutrition pellets; a ~3 g mixture of chocolate, banana, grape, and plain flavored pellets as a daily ration) and food restricted to approach 80-85% of their free-feeding body weight over the next 3 days before starting training on the neuroeconomic operant decision-making paradigm termed "Restaurant Row." Mice were weighed twice daily for the remainder of the experiment, before and after each Restaurant Row testing session. Mice were tested on Restaurant Row 7 days a week since performance on this task served as their sole source of food in this closed economy system. Mice foraged daily in a square maze for food rewards of varying cost (delays cued by tone pitch) and flavor (signaled by spatial context) while on a daily limited time budget (60 min). Each uniquely flavored "restaurant" was spatially fixed in the maze with patterns on the wall to signify the restaurant identity (chocolate: vertical stripes; banana: checkers; grape: triangles; plain, horizontal stripes). Rewards were delivered using a 3D printed automated pellet dispenser that was triggered by a computer running the behavioral task programmed in the ANY-Maze software made by the Stoelting Company. Behavioral events were triggered by spatial movements through the maze video-tracked by ANY-Maze. ELP USB camera with a Xenocam 1/2.7" 3.6 mm lens was used for video tracking. Restaurant Row testing took place in dim lighting conditions. The receptacle of the pellet dispenser also featured a custom-built trap door that would discard an uneaten pellet triggered upon exit from the wait zone if mice did not immediately consume food off of the pedestal. This prevented mice from hoarding rewards and forced animals to adhere to the structure of the task to make meaningful and intentional foraging decisions. Small wall mounted speakers (MakerHawk 3 Watt 8 Ohm Single Cavity Mini Speakers driven by a DROK 5W+5W Mini Amplifier Board PAM8406 DC 5V Dual Channel Class D) were fixed to the wall of each restaurant that played a 500 ms tone upon entry into the offer zone and repeated every s until either an enter or skip decision was made. The pitch of the tone varied depending on the randomly selected offer of that trial (1 s = 4,000 Hz and each second above that was an additional 387 Hz; e.g., 5 s = 5,548 Hz; 15 s = 9,418 Hz; 30 s = 15,223 Hz). Upon entry into the wait zone, the tones descended in a countdown fashion stepping down 387 Hz each second until a reward was earned and delivered. There is no penalty to quitting during the countdown other than the offer was rescinded and the mouse must advance to the next restaurant.