

## Prediction error in models of adaptive behavior

### Highlights

- The impact of predictive and nonpredictive relationships was assessed in rats
- Both relationships generated conditioned responding
- However, the nature and timing of responding was affected by these relationships
- The results are predicted by a recent model of associative learning, HeiDI

### Authors

Victor M. Navarro, Dominic M. Dwyer,  
Robert C. Honey

### Correspondence

Honey@cardiff.ac.uk

### In brief

Navarro et al. study the impact of predictive and nonpredictive relationships on conditioned responding in rats. The results contradict standard models of conditioning, which emphasize predictive relationships, but are predicted by the HeiDI model, which assumes that both relationships generate reciprocal associations that support responding.

Report

# Prediction error in models of adaptive behavior

Victor M. Navarro,<sup>1</sup> Dominic M. Dwyer,<sup>1</sup> and Robert C. Honey<sup>1,2,\*</sup>

<sup>1</sup>School of Psychology, Cardiff University, 70 Park Place, CF10 3AT Cardiff, UK

<sup>2</sup>Lead contact

\*Correspondence: [Honey@cardiff.ac.uk](mailto:Honey@cardiff.ac.uk)

<https://doi.org/10.1016/j.cub.2023.08.043>

## SUMMARY

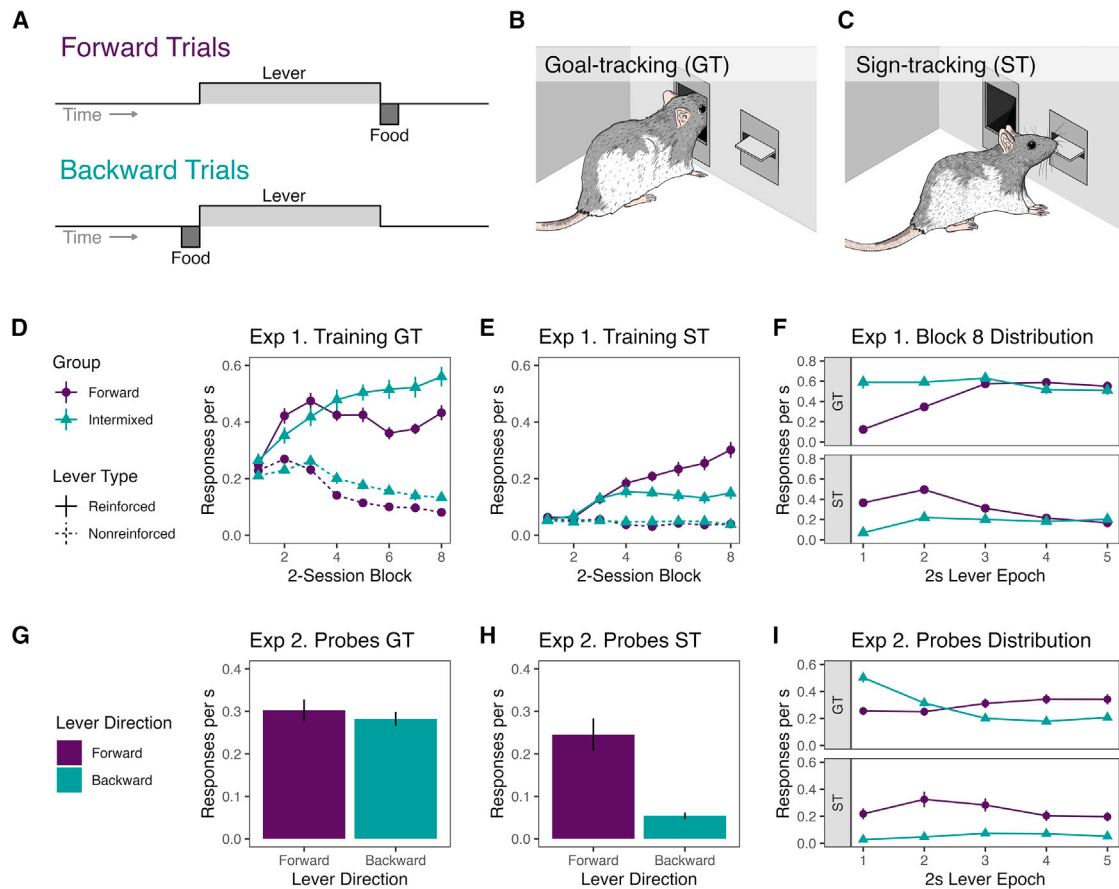
Pavlovian conditioning is evident in every species in which it has been assessed, and there is a consensus about its interpretation across behavioral,<sup>1,2</sup> brain,<sup>3–6</sup> and computational analyses<sup>7–11</sup>: conditioned behavior reflects the formation of a directional associative link from the memory of one stimulus (e.g., a visual stimulus) to another (e.g., food), with learning stopping when there is no error between the prediction generated by the visual stimulus and what happens next (e.g., food). This consensus fails to anticipate the results that we report here. In our experiments with rats, we find that arranging predictive (visual stimulus → food) and non-predictive (food → visual stimulus) relationships produces marked and sustained changes in conditioned behaviors when the visual stimulus is presented alone. Moreover, the type of relationship affects (1) the distribution of conditioned behaviors related to the properties of both food (called goal-tracking) and the visual stimulus (called sign-tracking) and (2) when in the visual stimulus, these two behaviors are evident. These results represent an impetus for a fundamental shift in how Pavlovian conditioning is interpreted: animals learn about the relationship between two stimuli irrespective of the order in which they are presented, but they exhibit this knowledge in different ways. This interpretation and our new results are captured by a recent model of Pavlovian conditioning,<sup>12,13</sup> HeiDI, and both are consistent with the need for animals to represent the fact that the impact of a cause (e.g., the ingestion of nutrients or the bite of a predator) can be felt *before* or *after* the cause has been perceived.

## RESULTS AND DISCUSSION

Standard theoretical models of Pavlovian conditioning<sup>7–11</sup> predict that while visual stimulus → food pairings result in the formation of a directional associative link from the memory of the visual stimulus to that of food, food → visual stimulus pairings result in the visual stimulus developing an inhibitory link with food: the two forms of trial are said to result in opposing forms of learning, with one enabling the presentation of the visual stimulus to excite a memory of food and the other enabling the visual stimulus to inhibit that memory. Certainly, when food is presented 10 s before a visual stimulus, it acquires such inhibitory properties,<sup>14</sup> which could derive from a temporary increase in the association between the experimental context and food, resulting in food being predicted but not occurring after the visual stimulus.<sup>12</sup> However, when food is presented just before the visual stimulus (i.e., they are temporally contiguous), the evidence for any form of learning is limited and equivocal. Many years ago, Mackintosh<sup>15</sup> concluded that “*All in all, there is little reason to accept the reality of backward conditioning*” (p. 60), by which he was referring to backward conditioning generating an excitatory link between the memories of the visual stimulus and food. In fact, the ongoing emphasis on predictive relationships has been fueled by the prevailing view that the behavioral sequelae of backward conditioning are weak and transient.<sup>16,17</sup> However, although the behavioral sequelae of forward conditioning have been investigated in

great detail, those of backward (excitatory) conditioning have not.

Unlike standard theoretical models, Asratyan<sup>18</sup> suggested that visual stimulus → food pairings result in the formation of reciprocal associations between the memories of the visual stimulus and food (i.e., visual stimulus ↔ food) and assumed that food → visual stimulus pairings should do likewise. On this basis, he argued that intermixing visual stimulus → food with food → visual stimulus pairings should produce complementary effects rather than the opposing (excitatory and inhibitory) effects envisaged by standard models.<sup>7–11</sup> This prediction is integral to a more formal model, HeiDI (How excitation and inhibition Determine Ideo-motion), in which reciprocal associations play a central role. HeiDI also includes performance rules that specify how the combined influence of the reciprocal associations affects the differing behavioral sequelae of Pavlovian conditioning across the duration of the visual stimulus: behaviors that reflect the nature of the visual stimulus (e.g., orienting) and food (e.g., approaching the site of food delivery).<sup>12,13</sup> We will return to this model later, but the general idea that learning involves the encoding of reciprocal relationships between the visual stimulus and food is also supported by recent neurobiological analyses.<sup>19</sup> Experiments 1 and 3 evaluated whether intermixing the two types of trial (forward and backward) has the predicted complementary behavioral effects or the opposing effects predicted by standard models. Experiment 2 contrasted the impact of forward and backward trials on the same measures.



**Figure 1. Forward and backward conditioning trials generate marked conditioned behavior**

(A) An autoshaping procedure was used to investigate the impact of forward and backward conditioning trials on the emergence of different conditioned behaviors. In this procedure, rats receive trials in which the temporary insertion of a lever into the experimental chamber is paired with the delivery of food.

(B and C) Autoshaping generates two types of conditioned behavior during the lever that can be easily measured in an automated fashion: approaching the well into which food is about to be delivered, called goal-tracking (GT; measured by the rate of food-well visits<sup>21</sup>), and orienting toward and interacting with the lever itself, called sign-tracking (ST; measured by the rate of lever presses<sup>20,22–24</sup>).

(D–F) Experiment 1 contrasted the impact of forward conditioning trials (lever 1 → food) with intermixed forward and backward conditioning trials (lever 1 → food and food → lever 1) on GT, ST, and their distribution across lever 1. A second lever (lever 2) was nonreinforced. Points represent mean response rates, and error bars represent SEM.

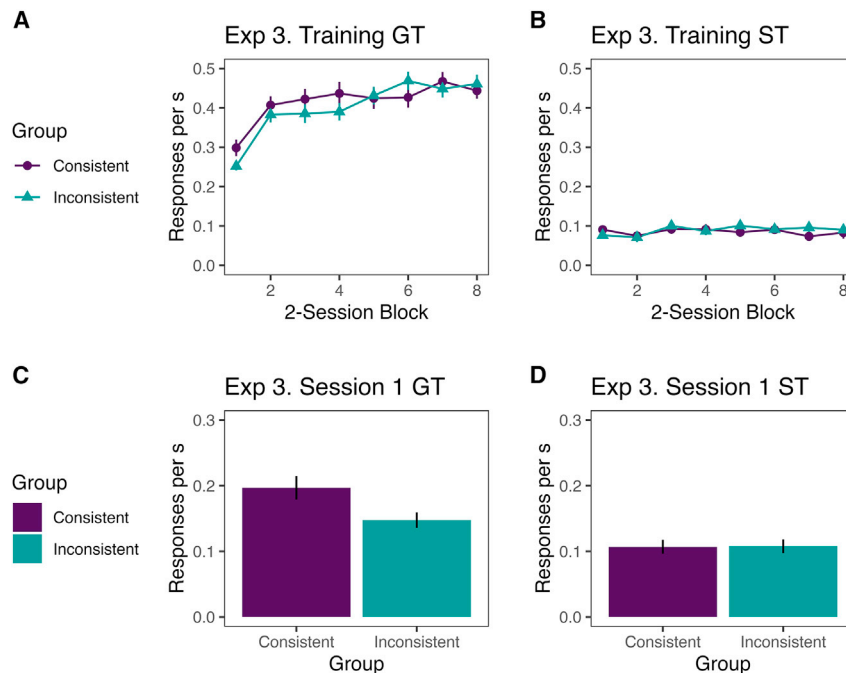
(G–I) In Experiment 2, all rats received forward conditioning trials with lever 1 and backward conditioning trials with lever 2, and the impact of these trials on GT, ST, and the distribution of these responses across levers 1 and 2 was assessed on nonreinforced probe trials. Points represent mean response rates, and error bars represent SEM.

### Intermixing forward and backward conditioning trials generates marked conditioned behavior

In Experiment 1, rats in group Forward ( $n = 32$ ) received forward conditioning trials on which the insertion of lever 1 (our visual stimulus) into an experimental chamber for 10 s was immediately followed by food (i.e., reinforced trials) and separate trials where the insertion of lever 2 for 10 s was not paired with food (i.e., nonreinforced trials; see [STAR Methods](#) for further details). This procedure generated goal-tracking (Figure 1D) and sign-tracking (Figure 1E) during lever 1, and during the final block of training (when performance was stable) goal-tracking was more vigorous during the final parts of lever 1 and sign-tracking was more apparent at the start of lever 1 (Figure 1F; replicating previous results<sup>20</sup>). There was little goal- or sign-tracking to lever 2, which was not paired with food (Figures 1D and 1E). In group Intermixed, half of the presentations of lever 1 were followed by

food and the remainder were preceded by food. Presentations of lever 2 were nonreinforced. Conditioned behaviors were assessed on the forward trials with lever 1 when consummatory responses generated by the presentation of food could not contaminate conditioned behaviors during the lever. There was at least as much goal-tracking in group Intermixed as in group Forward training, but less sign-tracking (Figures 1D and 1E). That is, the effects of the two types of trial appeared to be complementary for goal-tracking, but not for sign-tracking. The basis for these distinct effects is evident from an analysis of the final block of training: in group Intermixed, goal-tracking was vigorous across all epochs of lever 1 and sign-tracking was less vigorous across lever 1 (Figure 1F).

For all experiments, we estimated non-linear mixed-effects models using a Bayesian framework and performed inference on median posterior differences (MPDs) against a region of



**Figure 2. Interaction between forward and backward conditioning trials**

(A and B) The impact of consistent training (lever 1 → food 1 and food 1 → lever 1; lever 2 → food 2 and food 2 → lever 2) and inconsistent training (lever 1 → food 1 and food 2 → lever 1; lever 2 → food 2 and food 1 → lever 2) on goal-tracking (GT) and sign-tracking (ST) across 8 blocks of training. Points represent mean response rates, and error bars represent SEM.

(C and D) The impact of consistent and inconsistent training on GT and ST during the first training session. Bars denote means and error bars represent SEM.

practical equivalence (ROPE; see [STAR Methods](#)). Lever 1 provoked goal-tracking that was similar to that exhibited by group Forward (Figure 1D; MPD = 0.09, percentage of MPD in the region of practical equivalence [% in ROPE] = 0.07%). However, the same treatment led to weaker levels of sign-tracking (Figure 1E; MPD = 0.12, 0% in ROPE). Moreover, unlike the rats given only forward conditioning trials with lever 1, those given intermixed training showed similarly high levels of goal-tracking and low levels of sign-tracking across all epochs of the lever (Figure 1F). Indeed, goal-tracking during the last epoch was higher than during the first epoch for group forward (1F, MPD = 0.49, 0% in ROPE) but similar for group Intermixed (MPD = -0.07, 20.19% in ROPE); sign-tracking during the last epoch was lower than during the first epoch for group Forward (MPD = 0.21, 0% in ROPE), but similarly low for group Intermixed (MPD =  $2 \times 10^{-6}$ , 76.57% in ROPE). Again, there was little goal-tracking or sign-tracking to lever 2, which was not paired with food (Figures 1D and 1E); and in none of the experiments were there differences in the baseline levels of goal-tracking in the 10-s periods before lever presentations. The observation that backward conditioning trials change the nature and temporal distribution of conditioned behavior was replicated in Experiment 2; and this pattern of results is inconsistent with the idea that backward conditioning trials are simply *de facto* forward conditioning trials. Simulations of the results using HeiDI will be discussed once the results of Experiments 2 and 3 have been presented.

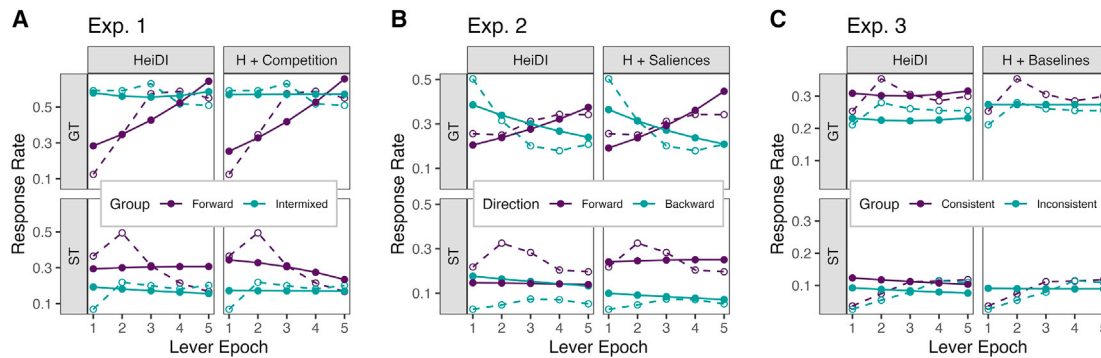
### Contrasting the effects of forward and backward conditioning trials

Experiment 2 directly contrasted the impact of forward and backward conditioning trials on goal-tracking and sign-tracking. All rats ( $n = 32$ ) received forward conditioning trials with one lever (e.g., lever 1 → food) intermixed with backward conditioning trials with a second lever (e.g., food → lever 2), and they received subsequent

probe trials in which levers 1 and 2 were presented in the absence of food, which avoids the immediate impact of food consumption on conditioned behaviors to lever 2. The probe trials revealed that backward trials, like forward trials, were highly effective in generating goal-tracking (Figure 1G) but were less effective in generating sign-tracking (Figure 1H). Indeed, the probe trials for the forward and backward levers provoked marked and similar goal-tracking (Figure 1G, MPD = 0.01, 24.35% in ROPE), but the forward lever elicited more sign-tracking than the backward lever (Figure 1H, MPD = 0.17, 0% in ROPE). The distribution of the two responses across the duration of the levers also differed. After forward trials with lever 1, goal-tracking was more evident during the late than the early epochs. Relative to the first epoch, goal-tracking during the last epoch was significantly higher for the forward lever (Figure 1I, MPD = 0.14, 0% in ROPE) but significantly lower for the backward lever (MPD = 0.35, 0% in ROPE). However, sign-tracking was stable across lever epochs (Figure 1I); relative to the first epoch, sign-tracking during the last epoch was not significantly different for either of the levers (MPD = 0.02, 46.77% in ROPE and MPD = 0.01, 71.37% in ROPE, for forward and backward levers, respectively).

### Interaction between forward and backward conditioning trials

Experiment 3 tested whether the effects of forward and backward trials interact. Rats received forward and backward trials in which there was either a consistent relationship between two levers and two foods (lever 1 → food 1 and food 1 → lever 1; lever 2 → food 2 and food 2 → lever 2;  $n = 32$ ) or there was an inconsistent relationship (lever 1 → food 1 and food 2 → lever 1; lever 2 → food 2 and food 1 → lever 2;  $n = 32$ ). We predicted that consistent training would result in a faster rate of learning than inconsistent training: for consistent training, backward and forward training trials with a given lever and food affect changes in the same associative structures (i.e., lever 1 ↔ food 1 and lever 2 ↔ food 2); whereas for inconsistent training, what is learned on backward trials will interfere with performance on forward trials and vice versa. In contrast, and for the reasons outlined in the context of Experiment 1, standard models of Pavlovian conditioning predict that what is learned on backward trials will be more likely to oppose what is learned on forward trials in rats given consistent training than in those given



**Figure 3. HeiDI fits to the experimental results**

(A–C) Model fits to experimental data across lever bins during the last block of training of Experiment 1 (A), the probe trials of Experiment 2 (B), and the first block of training of Experiment 3 (C). In each panel, closed symbols and solid lines denote the fits by the base HeiDI model (left column) and the expanded model with the lowest negative log likelihood (right column). The corresponding experimental results are shown as means in open symbols and dashed lines (GT, goal-tracking and ST, sign-tracking).

inconsistent training. However, note that in Experiment 3, unlike Experiment 1, predictions based on HeiDI and standard models rely on the rats being sensitive to the distinct properties of the two foods.

Inspection of Figure 2 revealed a rapid increase in goal-tracking and the dominance of goal-tracking over sign-tracking. This reflects the fact that both levers are always either preceded or succeeded by the presentation of food (1 or 2). It is also clear that during session 1 (Figure 2C), if not throughout training (Figure 2A: 2-session blocks), consistent training (as assessed on forward conditioning trials; MPD = 0.04, 95% highest posterior density of the MPD [0.002, 0.08], 2.29% in ROPE and MPD = 0.01, 34.19% in ROPE for session 1 rates and overall response rates, respectively). There was no effect of the nature of training on the low levels of sign-tracking at any point during training (Figure 2B: 2-session blocks; Figure 2D: session 1). The levels of goal-tracking and sign-tracking tended to rise across epochs of the lever, with the higher levels of goal-tracking in the group Consistent than in group Inconsistent across successive epochs on forward trials (not shown).

### Simulations of experimental results using the HeiDI model

HeiDI involves two principal departures from standard models which allow it to predict the results of Experiments 1–3: (1) the associative structures that are acquired and (2) the rules governing their behavioral expression. HeiDI assumes that reciprocal associations between stimuli develop, with the perceived intensity of the unconditioned stimulus (e.g., food) and conditioned stimulus (e.g., lever) determining the asymptotes for the lever→food and food→lever associations, respectively. The learning rules for these reciprocal associations involve a rationalization of the well-known Rescorla and Wagner rule.<sup>7</sup> To account for systematic variation in the timing of behaviors across a conditioned stimulus, HeiDI assumes that the perceived intensity of the lever declines across its presentation (e.g., due to short-term habituation<sup>20</sup>) and that the specific perceived intensity of the lever when food is delivered becomes linked to the

memory of food.<sup>13</sup> Under those assumptions, the conditioned lever intensity will be high on backward conditioning trials and low on forward conditioning trials, with the similarity of the remaining lever intensities to these two conditioned intensities providing a basis for generalization. Furthermore, the rules governing the behavioral expression of the reciprocal associations assume that when a lever is presented, the relative perceived intensities of the lever and the memory of food (that it retrieves) determine the distribution of what has been learned into behaviors related to the nature of the lever (e.g., sign-tracking) and food (e.g., goal-tracking). Other things being equal, when the lever has a perceived intensity that is higher than the perceived intensity of the memory of food then sign-tracking will dominate, but when the reverse is the case then goal-tracking will dominate.

Figure 3 depicts model fits to the data from Experiments 1–3, using the base HeiDI model<sup>12,13</sup> and additional versions of the model equipped with subsidiary mechanisms. For the base HeiDI model, we estimated the US intensity, the perceived intensity of the lever at its onset (i.e., the intensity conditioned with the food on backward pairings), and the rate and shape parameters of the power decay function that determined the decrease of lever intensity across its duration (and thus the final intensity conditioned with food on forward pairings). Three additional extensions were considered: (1) a model with additional baseline response rates for goal- and sign-tracking (H + Baselines); (2) a model with response competition (H + Competition), in which responses inhibit each other according to their relative strength; and (3) a model that estimated separate maximal lever (perceived) intensities or saliences for forward and backward trials (H + Saliences). Table 1 contains goodness-of-fit measures for each of those models (see STAR Methods).

The base model can capture the overall temporal dynamics of goal- and sign-tracking across the conditioned stimuli relatively well. In fact, metrics that penalize model complexity (AIC and BIC) nearly always positioned the base model as the most suitable to explain the data presented here (Table 1). The base model can account for each group in Experiment 1 separately (Figure 3A, bottom row), and its best-fitting parameters (not

**Table 1. Goodness-of-fit measures for HeiDI and its extensions**

Model	Exp. 1			Exp. 2			Exp. 3		
	NLL	AIC	BIC	NLL	AIC	BIC	NLL	AIC	BIC
HeiDI	604.33	1224.67*	1254.82*	600.51	1209.01	1226.86*	592.54	1193.07*	1210.92*
H + Baselines	605.20	1234.41	1279.63	600.36	1212.72	1239.49	592.46*	1196.91	1223.68
H + Saliences	605.93	1241.86	1298.39	599.03*	1208.07*	1230.38	592.85	1195.71	1218.02
H + Competition	603.59*	1227.19	1264.87	600.49	1210.98	1233.28	593.17	1196.34	1218.64

Note: NLL, negative log-likelihood; AIC, Akaike information criterion; BIC, Bayesian information criterion. For all measures, smaller values are better; and the smallest values for each metric/experiment combination are denoted by asterisks.

shown) suggest that food delivery on backward trials led to a general disruption in the processing of the lever.<sup>25–27</sup> Indeed, the base model predicts equal degrees of sign-tracking for forward and backward levers in Experiment 2 (Figure 3B, bottom left); but equipping it with lower lever saliencies for backward than forward conditioning trials results in closer fits (Figure 3B, bottom right). Notably, the base model provided a good fit to the different impacts of consistency on sign- and goal-tracking observed in session 1 of Experiment 3. A consistent difference between the models and the data is that the models do not capture the initial rise in sign-tracking. As we have noted elsewhere,<sup>20</sup> the low levels of sign-tracking (and goal-tracking) upon lever presentation are likely affected by the time taken for the rats to approach and contact the lever and food well.

## Conclusion

Our results and model represent a basis for a significant revision in how learning is understood within the model system of Pavlovian conditioning. They are consistent with the idea that reciprocal associative links (i.e., conditioned stimulus  $\rightleftharpoons$  unconditioned stimulus) develop during Pavlovian conditioning trials and these links impact behavior.<sup>12,13</sup> This idea has been implemented formally and provides a coherent explanation for a broad range of results that have proven resistant to analysis by standard models, based on a directional link between a conditioned and unconditioned stimulus.<sup>7–11</sup> More generally, the view that adaptive behavior reflects the development of error-driven reciprocal associations, combined with performance rules that capture hitherto ignored aspects of the nature of conditioned behavior, represents a clear impetus for further research. This analysis, coupled with our results, is consistent with the fact that stimuli presented *before or after* nutrients have been ingested, or a predator has attacked, provide useful information to guide future approach and avoidance behaviors.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS

- Regulations
- Apparatus
- Pre-training
- Conditioning procedures
- QUANTIFICATION AND STATISTICAL ANALYSIS
- Model fits

## ACKNOWLEDGMENTS

The research was supported by the BBSRC (UK; BB/T004339/1; PI: R.C.H.): “How knowledge affects behaviour in a model system.” We thank: Advanced Research Computing at Cardiff (ARCCA) for providing the computing resources necessary for the analyses presented here, A.F. Ilescu for help with conducting Experiment 1, and Leonnore Diaz for the illustrations depicted in Figure 1.

## AUTHOR CONTRIBUTIONS

V.M.N., D.M.D., and R.C.H. contributed to the ideas presented in this article, the design of the experiments, the analysis and modeling of the results, and to writing the paper. V.M.N. conducted Experiments 2 and 3.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: January 20, 2023

Revised: July 19, 2023

Accepted: August 15, 2023

Published: September 13, 2023

## REFERENCES

1. Rescorla, R.A. (1988). Pavlovian conditioning: It's not what you think it is. *Am. Psychol.* 43, 151–160.
2. Pavlov, I.P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex* (Oxford Univ. Press).
3. Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuji, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E.M., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature* 469, 53–57. <https://doi.org/10.1038/nature09588>.
4. Maes, E.J.P., Sharpe, M.J., Usypchuk, A.A., Lozzi, M., Chang, C.Y., Gardner, M.P.H., Schoenbaum, G., and Iordanova, M.D. (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. *Nat. Neurosci.* 23, 176–178. <https://doi.org/10.1038/s41593-019-0574-1>.
5. Mollick, J.A., Hazy, T.E., Krueger, K.A., Nair, A., Mackie, P., Herd, S.A., and O'Reilly, R.C. (2020). A systems-neuroscience model of phasic dopamine. *Psychol. Rev.* 127, 972–1021. <https://doi.org/10.1037/rev0000199>.

6. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>.
7. Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: Current research and theory*, A.H. Black, and W.F. Prokasy, eds. (Appleton-Century-Crofts), pp. 64–69.
8. Seymour, B., O’Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., Friston, K.J., and Frackowiak, R.S. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667. <https://doi.org/10.1038/nature02581>.
9. Sutton, R.S., and Barto, A.G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychol. Rev.* 88, 135–170. <https://doi.org/10.1037/0033-295X.88.2.135>.
10. Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* 82, 276–298.
11. Pearce, J.M., and Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532–552. <https://doi.org/10.1037/0033-295X.87.6.532>.
12. Honey, R.C., Dwyer, D.M., and Iliescu, A.F. (2020). HeiDI: A model for Pavlovian learning and performance with reciprocal associations. *Psychol. Rev.* 127, 829–852. <https://doi.org/10.1037/rev0000196>.
13. Honey, R.C., and Dwyer, D.M. (2022). Higher-order conditioning: A critical review and computational model. *Psychol. Rev.* 129, 1338–1357. <https://doi.org/10.1037/rev0000368>.
14. Delamater, A.R., LoLordo, V.M., and Sosa, W. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Learn. Behav.* 31, 393–402. <https://doi.org/10.3758/BF03196000>.
15. Mackintosh, N.J. (1974). *The psychology of animal learning* (Academic Press).
16. Heth, C.D. (1976). Simultaneous and backward fear conditioning as a function of number of CS-UCS pairings. *J. Exp. Psychol. Anim. Behav. Process.* 2, 117–129. <https://doi.org/10.1037/0097-7403.2.2.117>.
17. Matzel, L.D., Held, F.P., and Miller, R.R. (1988). Information and expression of simultaneous and backward associations: Implications for contiguity theory. *Learn. Motiv.* 19, 317–344. [https://doi.org/10.1016/0023-9690\(88\)90044-6](https://doi.org/10.1016/0023-9690(88)90044-6).
18. Asratyan, E.A. (1965). *Compensatory adaptations, reflex activity and the brain* (Pergamon Press).
19. Jeong, H., Taylor, A., Floeder, J.R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D.A., and Nambodiri, V.M.K. (2022). Mesolimbic dopamine release conveys causal associations. *Science* 378, eabq6740. <https://doi.org/10.1126/science.abq6740>.
20. Iliescu, A.F., Dwyer, D.M., and Honey, R.C. (2020). Individual differences in the nature of conditioned behavior across a conditioned stimulus: Adaptation and application of a model. *J. Exp. Psychol. Anim. Learn. Cogn.* 46, 460–469. <https://doi.org/10.1037/xan0000270>.
21. Boakes, R.A. (1977). Performance on learning to associate a stimulus with positive reinforcement. In *Operant-Pavlovian Interactions* (Routledge).
22. Davey, G.C., and Cleland, G.G. (1982). Topography of signal-centered behavior in the rat: Effects of deprivation state and reinforcer type. *J. Exp. Anal. Behav.* 38, 291–304. <https://doi.org/10.1901/jeab.1982.38-291>.
23. Hearst, E., and Jenkins, H.M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action* (Psychonomic Society).
24. Timberlake, W., Wahl, G., and King, D. (1982). Stimulus and response contingencies in the misbehavior of rats. *J. Exp. Psychol. Anim. Behav. Process.* 8, 62–85.
25. Rescorla, R.A. (1980). Simultaneous and successive associations in sensory preconditioning. *J. Exp. Psychol. Anim. Behav. Process.* 6, 207–216. <https://doi.org/10.1037/0097-7403.6.3.207>.
26. Sutherland, N.S., and Mackintosh, N.J. (1971). *Mechanisms of Animal Discrimination Learning* (Academic Press).
27. Thomas, D.R. (1970). Stimulus selection, attention, and related matters. In *Current issues in animal learning*, J.H. Reynierse, ed. (University of Nebraska Press).
28. Kruschke, J.K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Adv. Methods Pract. Psychol. Sci.* 1, 270–280. <https://doi.org/10.1177/2515245918771304>.
29. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. (L. Erlbaum Associates).
30. R Development Core Team (2021). R: A language and environment for statistical computing. <https://www.r-project.org/>.
31. Bürkner, P.-C., Gabry, J., Weber, S., Johnson, A., Modrak, M., Badr, H.S., Weber, F., Ben-Shachar, M.S., and Rabel, H. (2022). brms: Bayesian regression models using “Stan”. <https://cran.r-project.org/web/packages/brms/>.
32. Kay, M., and Mastny, T. (2022). tidybayes: Tidy data and “geoms” for Bayesian models. <https://cran.r-project.org/web/packages/tidybayes>.
33. Makowski, D., Lüdtke, D., Ben-Shachar, M.S., Patil, I., Wilson, M.D., Wiernik, B.M., Bürkner, P.-C., Mahr, T., Singmann, H., Gronau, Q.F., et al. (2022). bayestestR: Understand and describe Bayesian models and posterior distributions. <https://cran.r-project.org/web/packages/bayestestR/>.
34. Lenth, R.V., Bürkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., and Singmann, H. (2022). emmeans: Estimated marginal means, aka least-squares means. <https://cran.r-project.org/web/packages/emmeans/>.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Organisms/strains		
Male Lister Hooded rats	Envigo (Blackthorn, UK)	N/A
Software and algorithms		
R programming language (version 4.1.2)	<a href="https://www.r-project.org/">https://www.r-project.org/</a>	N/A
brms: Bayesian regression models using “Stan” (version 2.17.0)	<a href="https://cran.r-project.org/web/packages/brms/">https://cran.r-project.org/web/packages/brms/</a>	N/A
BayestestR: Understand and describe Bayesian models and posterior distributions (version 0.11.5)	<a href="https://cran.r-project.org/web/packages/bayestestR/">https://cran.r-project.org/web/packages/bayestestR/</a>	N/A
emmeans: Estimated marginal means, aka least-squares means (version 1.7.3)	<a href="https://cran.r-project.org/web/packages/emmeans/">https://cran.r-project.org/web/packages/emmeans/</a>	N/A
calmr: Canonical associative learning models in R (version 0.3.0)	<a href="https://github.com/victor-navarro/calmr">https://github.com/victor-navarro/calmr</a>	N/A
Deposited data		
All data and code used in the preparation of this manuscript	OSF	<a href="https://doi.org/10.17605/osf.io/mprka">https://doi.org/10.17605/osf.io/mprka</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further inquiries not covered by those materials should be directed to Robert C. Honey ([Honey@cardiff.ac.uk](mailto:Honey@cardiff.ac.uk)).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

All data and code used in the preparation of this manuscript are available at the project’s OSF repository (<https://osf.io/mprka/>).

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of 160 naïve adult male Lister Hooded rats (mean *ad libitum* weight = 334 g; range = 260–410 g; supplied by Envigo, Blackthorn, UK) were used: Experiment 1 ( $n = 64$ ), Experiment 2 ( $n = 32$ ), and Experiment 3 ( $n = 64$ ). Rats were housed in groups (range: 2–4) in standard cages in a vivarium where there was a 12-h/12-h light/dark cycle (lights on at 7 a.m.). They were maintained at between 85% and 95% of their *ad libitum* weights by restricting their access to food until the end of each day. Access to food was given in their home cages, where they also had continuous access to water. Subjects were randomly assigned to each group before the beginning of each experiment.

### METHOD DETAILS

#### Regulations

The research was conducted following Home Office regulations under the Animal (Scientific Procedures) Act 1986. Approval of the study protocol was granted to Dominic M. Dwyer (Affect and Cognition in Rodents. PPL number: PP3468526).

#### Apparatus

Sixteen identical conditioning boxes (30 × 24 × 21 cm: H×W×D; Med Associates, Fairfax, VT, USA) were used. Each box was placed in a sound-attenuating shell incorporating a ventilation fan, which maintained the background noise at 68 dB(A). The side walls of the boxes were made from aluminium and the front, back and ceiling were clear acrylic. The floors were formed from 19 steel rods (4.8 mm diameter, 16 mm apart) and were positioned above a stainless-steel tray. Food pellets (45 mg; LabDiet, St. Louis, MO, USA) and an 8% sucrose solution were delivered to a recessed food well (aperture: 5.3 × 5.3 cm) positioned at floor level in the center of the left wall. Infrared detectors across the entrance to the food well allowed the presence of the rat to be recorded automatically.



When the detector was interrupted (e.g., when a rat's snout entered the food well) a single response was registered. Two retractable levers ( $4.5 \times 1.8 \times 0.2$  cm) were located on either side of the food well, one 3 cm to the left and one 3 cm to the right of the food well, and positioned at a height of 4.6 cm and 1.5 cm from the edge of the wall. Depression of the lever by 4mm from its horizontal resting position was recorded as a response. MED-PC IV software (Med Associates, Fairfax, VT, USA) controlled the insertion of levers, delivery of food pellets, and recorded food well entries and lever presses.

### Pre-training

In two 21-min sessions, rats were first trained to retrieve the unconditioned stimulus (food) or unconditioned stimuli (food and sucrose) used in the conditioning procedures. In each session, the rats received 20 unconditioned stimulus presentations (mean inter-trial interval = 60 s; sampled uniformly from 40, 50, 60, 70, and 80 s). In Experiment 1, standard food pellets were presented, whereas in Experiments 2 and 3, pellets were used alongside sucrose.

### Conditioning procedures

There were 10 presentations of each lever during each 20-min session, which were scheduled in the same way as pre-training. In all experiments, forward trials involved the presentation of a lever for 10 s, upon which food was delivered, whereas backward trials involved the delivery of food, a delay (1.0 or 1.5 s for pellet and sucrose, respectively), and then the presentation of a lever for 10 s. Non-reinforced trials and probe trials involved the presentation of a lever for 10 s. Trials were pseudo-randomized such that no trial type occurred more than 2 times in succession. The identities of each lever were counterbalanced across rats, as were and the two food types (pellets and sucrose) used in Experiments 2 and 3.

Experiment 1. Rats ( $n = 64$ ) received 16 sessions of training. In each session, group Forward ( $n = 32$ ) received 10 forward trials involving lever 1 and 10 non-reinforced trials of lever 2, while rats in group Intermixed ( $n = 32$ ) received 5 forward trials and 5 backward trials involving lever 1, and 10 non-reinforced presentations of lever 2.

Experiment 2. Rats ( $n = 32$ ) received 16 training sessions. These sessions employed the same procedure as Experiment 1, and each contained 10 forward trials with lever 1 and 10 backward trials with lever 2. For half of the rats, both levers were paired with the same outcome (a food pellet or sucrose), whereas for the other half, one lever was followed by a food pellet and the other by sucrose. Whether the same or different outcomes were paired with the two levers did not affect the pattern of results, with the exception that when two outcomes were used the overall level of behavior tended to be higher. To simplify the analysis, this factor was omitted in the presentation and analysis of the results. After training, rats received 2, 5-day test blocks in which days 1, 3, and 5 were training sessions (as described above), and days 2 and 4 were test sessions. In test sessions, the first 8 trials were forward and backward warm-up trials (as described for training; 4 trials with each lever). The remaining 12 trials were organized in blocks, each containing 1 training trial and 1 probe trial per lever (for a total of 4 trials per block). The probe trials were pooled across cycles for presentation and analysis because the results were highly consistent across the cycles of testing and retraining.

Experiment 3. Rats in group consistent ( $n = 32$ ) received 16 sessions of training in which the forward and backward trials with lever 1 involved food 1 and the trials of lever 2 involved food 2 (i.e., lever 1  $\rightarrow$  food 1 and food 1  $\rightarrow$  lever 1; lever 2  $\rightarrow$  food 2 and food 2  $\rightarrow$  lever 2). Rats in group inconsistent ( $n = 32$ ) received training trials in which the food paired with a given lever differed on forward and backward trials (i.e., lever 1  $\rightarrow$  food 1 and food 2  $\rightarrow$  lever 1; lever 2  $\rightarrow$  food 2 and food 1  $\rightarrow$  lever 2).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis was done via mixed-effects models estimated under a Bayesian framework. We chose mixed-effects models because all of our experiments involved within-subject comparisons, and thus much group-level parameter shrinkage was to be gained from capturing variance due to individual differences. We chose a Bayesian framework, because of its capacity to quantify uncertainty and its inherent ability in quantifying the evidence in favor of experimentally important null hypotheses. The response rates for nose pokes and lever presses were assessed separately for each experiment, via nonlinear, hurdle lognormal models. The nonlinear portion of the model estimated the rate of responding via  $rate \sim asym + (r0 - asym) * exp(-exp(lrc) * x)$ , where the  $exp$  function denotes exponentiation,  $asym$  is the asymptotic rate,  $r0$  is the rate when  $x = 0$ ,  $lrc$  is the natural logarithm of the rate constant, and  $x$  is the longitudinal variable used in the model (block or epoch). For analyses focusing on overall means, a simpler linear model was used. The linear model estimated the rate of responding via  $rate \sim X\beta$  where  $X$  is the design matrix and  $\beta$  is a vector containing its corresponding parameters. In parallel to the estimation of response rates, the hurdle part of the hurdle lognormal models estimated the probability of no responding. In all analyses, we regressed rates of responding as a function of fixed and random effects, but neither its dispersion nor hurdle to reduce model complexity. Statistical inference was performed based on the posterior distributions of median differences, accepting or rejecting these differences via practical equivalence tests using an HD + ROPE criterion.<sup>28</sup> This type of equivalence test computes the percentage of the posterior distribution that falls within a region of practical equivalence (ROPE). For the present application, we used a default ROPE in the range of  $-0.1\sigma$  to  $+0.1\sigma$  (where  $\sigma$  is the standard deviation of the response rates estimated by each model). Such a ROPE range represents negligible differences in the positive and negative directions.<sup>29</sup> Differences were deemed true null differences if the 95% central posterior distribution fell within the ROPE, or true significant differences if that same portion of the posterior distribution fell outside the ROPE. The test was deemed inconclusive if neither of the previous cases was true. All analyses were carried out in R,<sup>30</sup> using packages `brms`<sup>31</sup> (version 2.17.0), `tidybayes`<sup>32</sup> (version 3.0.2), `bayestestR`<sup>33</sup> (version 0.11.5), and `emmeans`<sup>34</sup> (version 1.7.3). All models were estimated by running 8 independent chains

of 4000 iterations each (1000 iterations were used as a warmup and were not included in the final posterior distribution). See OSF site (Data availability) for additional information on the approach, the exact model specification for each analysis and the full results. All models and R scripts used are available on the same site.

### Model fits

The implementation of the HeiDI model involves an adaptation process across the lever duration, described by a decay function of the form:  $\alpha_i^e = (\alpha_i^{e-1} - \alpha_{min}) \times (1 - \lambda)^\rho + \alpha_{min}$ , where  $\alpha_i^e$  is the perceived intensity or salience of the  $i$ th lever on epoch  $e$  (for  $e = 1$ , the maximal salience the lever takes),  $\alpha_{min}$  is the minimum salience lever can take (here, set to 0),  $\lambda$  is the rate of decay, and  $\rho$  is a parameter that determines the shape of the decay function. This adaptation process allows the model to provide an analysis of the distribution of different types of responses (here sign-tracking and goal-tracking) across the duration of the lever.<sup>20</sup> HeiDI assumes that reciprocal associations between stimuli (e.g., lever  $\rightarrow$  food and food  $\rightarrow$  lever associations) update according to a global error correction rule. For example, on trial  $n$  the association between the  $i$ th CS (conditioned stimulus; e.g., a lever) and the US (unconditioned stimulus; e.g., food) on epoch  $e$  updates according to:  $V_{i \rightarrow US}^{e,n} = V_{i \rightarrow US}^{e,n-1} + \alpha_i^e (\alpha_{US}^e - \sum_j V_{j \rightarrow US}^{e,n-1})$ , where  $\alpha_{US}^e$  is the perceived intensity or salience of the US on epoch  $e$ , and the set  $J$  contains all stimuli present on epoch  $e$ . The same equation can be used to update the association between the US and the CS (or CSs, if more than one CS is present at the time of US delivery). Furthermore, the model proposes that reciprocal associations combine into stimulus ensembles via direct and indirect pathways, with each weighted according to generalization between different perceived intensities of a given lever. Ignoring the trial notation for simplicity, the strength of the direct pathway between the  $i$ th lever and the US on epoch  $e$  is given by:  $O_{i \rightarrow US}^e = \sum_m S(\alpha_i^e, \alpha_i^m) \times V_{i \rightarrow US}^m (1 + V_{US \rightarrow i}^m)$ , where the set  $M$  contains all epochs of lever presentation (including  $e$ ) and the function  $S$  is a similarity function of the form  $S(e, m) = \frac{e}{e+|e-m|} \times \frac{m}{m+|e-m|}$ . The strength of the indirect pathway between the  $i$ th lever and the US on epoch  $e$  is given by:  $H_{i \rightarrow US}^e = \sum_m S(\alpha_i^e, \alpha_i^m) \times \sum_k S(\alpha_k^e, \alpha_k^m) V_{i \rightarrow k}^m O_{k \rightarrow US}^m$ , where the set  $K$  contains all absent stimuli,  $\alpha_k^e$  is the nominal salience of the  $k$ th stimulus if it was present on epoch  $e$ , and  $\alpha_k^e$  is the similarity-weighted activation of the  $k$ th stimulus on epoch  $e$ ,  $\alpha_k^e = \sum_j \sum_m S(\alpha_j^e, \alpha_j^m) V_{j \rightarrow k}^m$ . Finally, the net ensemble involving the US on epoch  $e$ ,  $A_{US}^e = \sum_j (O_{j \rightarrow US}^e + H_{j \rightarrow US}^e)$ , is distributed into stimulus-oriented responses according to the relative activation of all stimuli. The strength of responding supported by the  $x$ th stimulus (e.g., CS- or US-oriented responding) on epoch  $e$  is given by:  $R_x^e = A_{US}^e \times \frac{\theta_x^e}{\sum_x \theta_x^e}$ , where  $X$  is the set containing all experimental stimuli, and  $\theta_x^e$  is the nominal or retrieved salience of the  $x$ th stimulus on epoch  $e$ . If stimulus  $x$  is present ( $x \in J$ ), then  $\theta_x^e = \alpha_x^e$ , however, if stimulus  $x$  is absent ( $x \in K$ ), then  $\theta_x^e = \alpha_x^e$ , as described above.

The extensions of the base HeiDI model capture some plausible subsidiary mechanisms: response competition, baseline differences in the different types of response, and differences in lever salience generated on forward and backward training trials. The H + Competition model proposes that stimulus-oriented responses will inhibit each other according to their relative strength. In that model, the final responding supported by the  $i$ th stimulus on epoch  $e$  is given by:  $R_i^e = R_i^e - \sum_{x \neq i} \kappa R_x^e$ , where  $\kappa$  is a parameter between 0 and 1 governing the intensity of response competition. The H + Baselines model adds a stimulus-specific constant to the response rates,  $R_i^e = R_i^e + c_i$ . In the present application, we estimated a constant for US-oriented responses and a constant for lever-oriented responses shared among all lever stimuli. Finally, the H + Saliences model estimated separate lever saliences for forward and backward conditioned trials. The data shown in [Figure 3](#) used the salience value for forward trials (when food is not delivered until after the lever offset).

The parameters were estimated for each experiment separately (and for all rats within an experiment simultaneously, except for Experiment 1) via MLE on model residuals using a genetic algorithm run over 300 populations per model/experiment combination (convergence into stable estimates was obtained much earlier). The upper and lower limits for all parameters were set between 0 and 1, except for the limits for the shape parameter for the power decay function, which was set between 0 and 5.