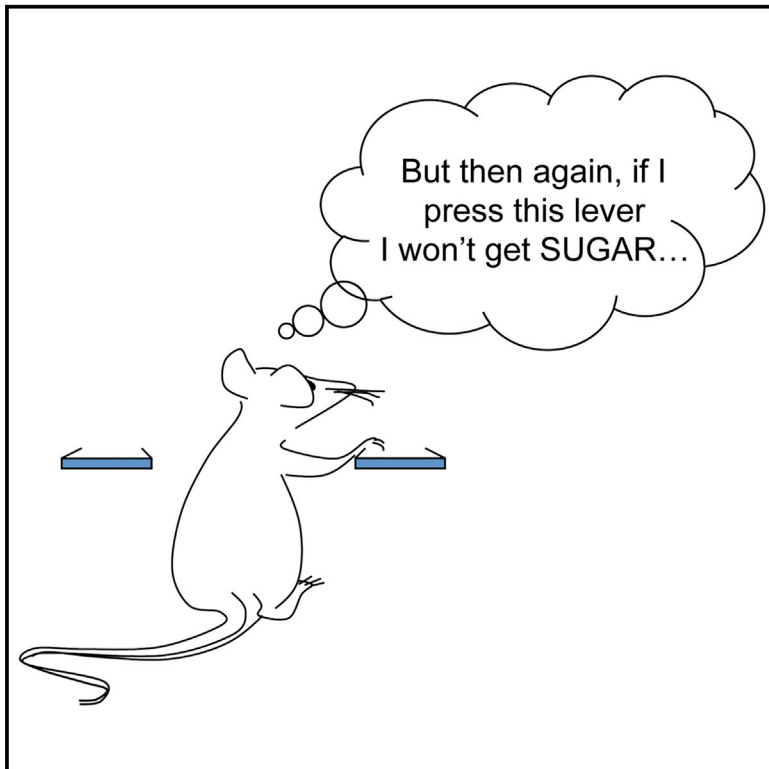# Current Biology

Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats

Graphical Abstract

Authors
Vincent Laurent, Bernard W. Balleine

Correspondence
bernard.balleine@sydney.edu.au

In Brief
Counterfactual reasoning has been argued to provide an adaptive advantage for human action. Here, Laurent and Balleine show that rats can reason counterfactually. They demonstrate that rats can encode consequences that their actions do not produce and can use that information to choose between alternative actions.

Highlights

- Rats encode the outcomes their actions produce and those they do not produce

- Positive outcome predictions guide choice to actions that earn that outcome

- Negative outcome predictions guide choice to actions that do not earn the outcome

- Rats can use counterfactual reasoning about outcome predictions to guide choice

CellPress

Factual and Counterfactual Action-Outcome Mappings Control Choice between Goal-Directed Actions in Rats

Vincent Laurent[1] and Bernard W. Balleine[2,*]
[1]School of Psychology, The University of Sydney, Sydney, NSW 2006, Australia
[2]Behavioural Neuroscience Laboratory, Brain and Mind Research Institute, The University of Sydney, Sydney, NSW 2006, Australia
*Correspondence: bernard.balleine@sydney.edu.au
http://dx.doi.org/10.1016/j.cub.2015.02.044

SUMMARY

The capacity to extract causal knowledge from the environment allows us to predict future events and to use those predictions to decide on a course of action [1]. Although evidence of such causal reasoning has long been described [2], recent evidence suggests that using predictive knowledge to guide decision-making in this way is predicated on reasoning about causes in two quite distinct ways: choosing an action can be based on the interaction between predictive information and the consequences of that action, or, alternatively, actions can be selected based on the consequences that they do not produce [3–5]. The latter counterfactual reasoning is highly adaptive because it allows us to use information about both present and absent events to guide decision-making [6, 7]. Nevertheless, although there is now evidence to suggest that animals other than humans, including rats and birds, can engage in causal reasoning of one kind or another [8–10], there is currently no evidence that they use counterfactual reasoning to guide choice. To assess this question, we gave rats the opportunity to learn new action-outcome relationships, after which we probed the structure of this learning by presenting excitatory and inhibitory cues predicting that the specific outcomes of their actions would either occur or would not occur. Whereas the excitors biased choice toward the action delivering the predicted outcome, the inhibitory cues selectively elevated actions predicting the absence of the inhibited outcome, suggesting that rats encoded the counterfactual action-outcome mappings and were able to use them to guide choice.

RESULTS

To assess whether evidence of counterfactual reasoning can be found in animals, we took advantage of a paradigm developed in rats to assess the influence of predictive learning on choice between actions that earn otherwise similarly valued outcomes. This paradigm, called Pavlovian-instrumental transfer, typically reveals that an event predictive of a particular outcome biases choice toward actions that earn the same outcome [11]. Although this finding suggests that rats can use predictive information to guide choice, it does not provide any direct evidence of counterfactual reasoning; whereas considerable evidence suggests that rats can encode specific action-outcome relationships [12, 13], there is currently no evidence that they encode the outcomes that an action does not produce. To ask this question in the context of Pavlovian-instrumental transfer, we framed it in a different way: in addition to positive predictions, we provided rats with information that a specific outcome would not occur and assessed how they modified their choice of action in the face of this information. Research in animal learning has found essentially three ways to establish a stimulus as a negative predictor. The first is by presenting the stimulus whenever a predicted outcome is omitted, called a feature-negative design [14, 15]; the second is by presenting the predictive stimulus whenever there is a reduction in the value of the expected outcome, called overexpectation [16]; and, the third is by reversing the usual causal order of stimulus and outcome, i.e., by backward conditioning [17]. In the current study, we used all three methods and assessed the effects of both the positive and the negative predictions on choice.

If, in the course of learning about actions and their various outcomes, i.e., during instrumental conditioning, rats are sensitive to counterfactual information, then we anticipate that they will learn not only what outcome is earned by performing a specific action but also what outcomes that action does not earn. The critical question is what the rats do when given information about the absence of a specific outcome. If they do not reason counterfactually, then such information should do no more than inhibit actions associated with the absent outcome. If, however, rats can reason counterfactually, then, in the presence of information about the absence of a specific outcome, they should be expected to increase their performance of actions that they have learned predict the absence of that outcome.

Experiment 1: Positive and Negative Predictions Derived from Feature-Negative Learning Exert Opposing Effects on Choice

Experiment 1 used a feature-negative or conditioned inhibition design to establish unique positive and negative predictors of two distinct outcomes and then evaluated their influence on choice between actions. The design is presented in Figure 1A

CrossMark
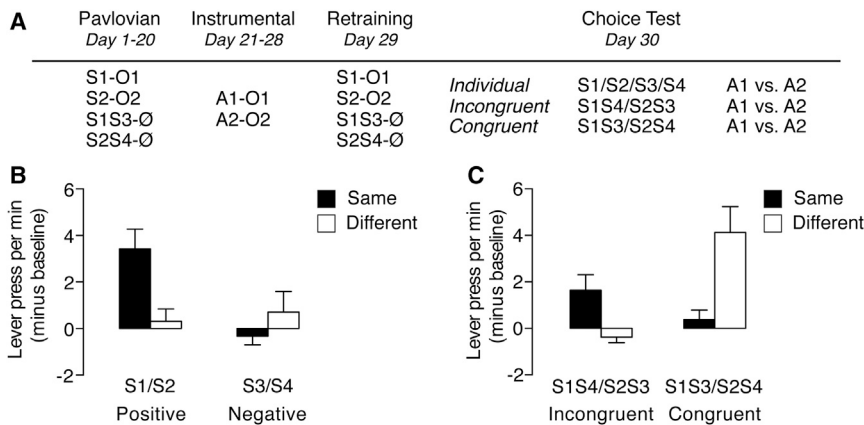
Figure 1. The Effect of Negative Predictions on Choice

(A) Design of the experiment using a feature-negative, conditioned inhibition procedure to generate negative predictors (n = 24). Abbreviations are as follows: S1/S2, noise and clicker stimuli (counterbalanced); S3/S4, house-light and tone stimuli (counterbalanced); O1/O2, food pellet and sugar outcomes (counterbalanced); Ø, no outcome; A1/A2, instrumental lever press actions on a left lever and right lever (counterbalanced).

(B) Performance during the various predictive stimuli was assessed relative to performance in the absence of the predictors by subtracting this latter baseline activity to reveal the net effect of the predictive cues on choice and the vigor of action selection. Positive (S1/S2) and negative (S3/S4) predictors had a distinct influence on choice between actions ($F_{(1,23)}$ = 7.17, p < 0.05). The positive predictors biased choice toward the action (i.e., Same) with which they shared a common outcome ($F_{(1,23)}$ = 10.77, p < 0.05) and elevated the performance of that action above baseline, whereas the negative predictors failed to trigger any significant preference between the two actions (F < 1.5).
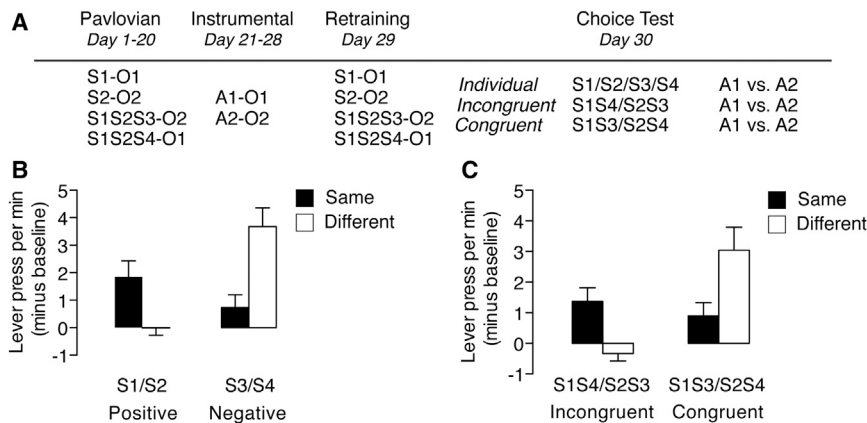
(C) Incongruent (S1S4/S2S3) and congruent (S1S3/S2S4) compounds exerted an opposite effect on choice ($F_{(1,23)}$ = 25.16, p < 0.05). The former increased responding on the Same action, which previously earned the outcome signaled by the positive predictor embedded in the compound ($F_{(1,23)}$ = 12.03, p < 0.05). In contrast, the congruent compound biased choice away from the Same action toward the Different action and increased the performance of that action above baseline, which was associated with the absence of the outcome that the compound stimuli signaled would not occur ($F_{(1,23)}$ = 9.71, p < 0.05). Error bars denote ±1 SEM.

and had three stages. The first stage involved Pavlovian training during which two stimuli, S1 and S2, were trained as Pavlovian excitors through pairings with the two distinct food outcomes, O1 and O2. Another two stimuli, S3 and S4, were trained as conditioned inhibitors of O1 and O2, respectively. This was achieved by repeatedly presenting two compounds, S1S3 and S2S4, in the absence of either outcome. As a consequence, we hoped to establish S3 as a predictor of "no O1" (S3 was presented whenever the S1 prediction of O1 was disconfirmed) and, for similar reasons, to establish S4 as a predictor of "no O2." The rats were then trained to perform the two lever press actions, A1 and A2, to deliver the two outcomes; i.e., A1 earned O1, whereas A2 earned O2, such that, by implication, A1 did not earn O2, and A2 did not earn O1. To establish whether rats encoded these counterfactual relationships and could use them to guide choice, we conducted a single test in which the levers were available but pressing them did not result in outcome delivery. This allowed us to assess the effects of the stimuli on choice in the absence of specific feedback. The positive and negative predictors were presented either alone or in compound. The compound trials were critical to establish the role of counterfactual reasoning because, in these trials, information was presented either congruent with training—where the negative predictor was congruent with the positive prediction, i.e., S1S3 and S2S4—or incongruent with training, such that the negative predictor was incongruent with the positive prediction, i.e., S1S4 and S2S3. Thus, in congruent trials, the positive and negative predictions related to the same outcome, e.g., S1 predicting O1 with S3 predicting O1 would be omitted, whereas in incongruent trials, they related to different outcomes, e.g., S1 predicting O1 and S4 predicting O2 would be omitted, leaving a net positive prediction for O1.

Training data from the first two phases are presented in Figure S1. The data of most interest, those from the transfer test, are shown in Figures 1B and 1C. They are plotted as the mean number of lever presses per minute when the stimulus or the compound predicted the delivery, or the absence, of the same outcome of the action (Same) and when the stimulus or the compound predicted the delivery, or the absence, of the different outcome from the action (Different). Thus, A1 was identified as "Same" and A2 as "Different" in the presence of S1, S3, or S1S3. Conversely, A2 was labeled as "Same" and A1 as "Different" during S2, S4, or S2S4. Further, baseline responding was subtracted from these rates of responding in order to reveal the net increase in choice performance; i.e., the net effect of the predictive stimuli. Baseline was defined as the mean number of lever presses per minute on both actions when the stimuli or compounds were absent (see Supplemental Information for more details). Consistent with the literature, predictive information about O1 and O2 derived from S1 and S2 was found to bias performance toward the action that previously predicted the same outcome, i.e., A1 and A2, respectively (Figure 1B). Although S3 and S4 alone produced only a modest bias (Figure 1B), the critical question was how the rats would respond during the congruent versus the incongruent compounds. As both compounds predicted the absence of any subsequent outcome, it is possible they would inhibit responding on both levers. In fact, whereas the incongruent compounds biased choice in much the same manner as S1 and S2 alone, the congruent compounds not only biased choice in the opposite direction, they elevated the performance of that action above baseline, providing evidence that the rats not only encoded the counterfactual relationships but were able to use these relationships flexibly to alter their action selection on test (Figure 1C). Rather than inhibiting both actions, rats elevated their performance on the lever not delivering the outcome the compound predicted would be omitted. It is the complete reversal of S1 and S2's effect on choice that was induced by the congruent compound that provides the strongest evidence the rats engaged in counterfactual reasoning; not only were they

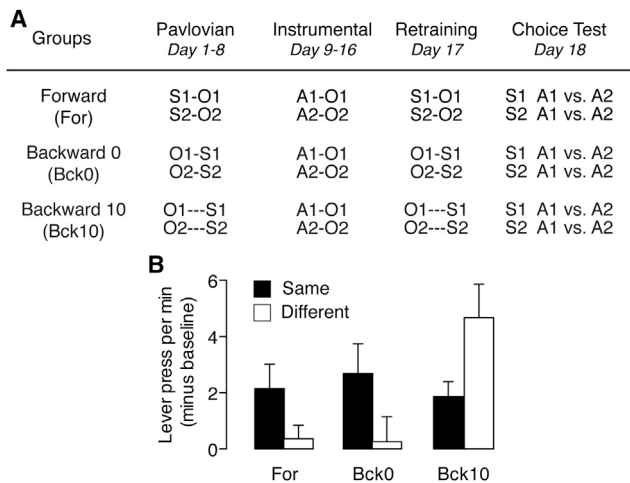**Figure 2. Negative Predictions Induced by Overexpectation**

(A) The overexpectation procedure used to generate negative predictors (n = 32). Abbreviations are the same as those in the legend of Figure 1.

(B) Positive (S1/S2) and negative (S3/S4) predictors exerted the opposite influence on choice between actions ($F_{(1,31)} = 20.73$, $p < 0.05$). The positive predictors biased choice toward the action with which they shared a common outcome (i.e., Same) ($F_{(1,31)} = 9.54$, $p < 0.05$), whereas responding on the Same action remained low in the presence of the negative predictors. In contrast, the latter stimuli directed choice toward the Different action, which was associated with the absence of the outcome that the stimuli also signaled would not occur ($F_{(1,31)} = 13.30$, $p < 0.05$).

(C) Incongruent (S1S4/S2S3) and congruent (S1S3/S2S4) compounds also had opposing effects on choice ($F_{(1,31)} = 11.37$, $p < 0.05$). Incongruent compounds promoted responding on the Same action, which previously delivered the outcome signaled by the positive predictor embedded in the compound ($F_{(1,31)} = 5.64$, $p < 0.05$). In contrast, the congruent compound both biased choice away from the Same action and elevated the performance of Different action ($F_{(1,31)} = 11.87$, $p < 0.05$), which was associated with the absence of the outcome that the compound stimuli signaled would not occur.
Error bars denote ±1 SEM.

sensitive to the specific negative predictors, but they could clearly use the information they conveyed (i.e., "no O1" or "no O2") to perform the alternative action.

Experiment 2: Overexpectation Produces Specific Negative Predictions that Reverse Choice

In addition to demonstrating counterfactual reasoning, the results from our first study showed that predictions about the absence of a specific outcome can be generated when that outcome is unexpectedly omitted. We used a similar logic in our next experiment to generate a specific negative predictor through overexpectation [16]. The design is presented in Figure 2A and, to the best of our knowledge, constitutes the first attempt to demonstrate an outcome-specific overexpectation effect. During the Pavlovian conditioning phase, a new set of naive rats again learned that S1 and S2 predicted O1 and O2, respectively. However, in this study, rats were also exposed to two compounds each composed of the two excitors (S1 and S2) plus a novel stimulus, either S3 or S4. One compound, S1S2S3, was paired with O2, whereas S1S2S4 was paired with O1. We assumed that combined exposure to S1 and S2 would predict the delivery of both O1 and O2. However, as O1 was omitted in the presence of S3 and O2 in the presence of S4, we hoped that these stimuli would become negative predictors of O1 and O2, respectively.

Using these stimuli, we applied the same logic as the previous study to assess the influence of positive and negative predictors of the two outcomes on choice between the two lever press actions. The rats were given lever press training as previously described and then tested using the two positive predictors, S1 and S2, and the same congruent and incongruent compounds used in the previous study (i.e., S1S3 and S2S4 versus S1S4 and S2S3) to provide information about the presentation or omission of O1 and O2 and assess the effect of this information on choice. The results from the training phases are presented in Figure S2. The results from the test are presented in Figures 2B and 2C in the manner described previously (see Supplemental Information for more details). We were, again, able to

replicate the bias in choice toward the action that previously earned the same outcome as that predicted by the positive predictors, S1 and S2, when these were presented alone (Figure 2B). Similarly, the same bias was also observed in the presence of the incongruent compounds (Figure 2C). In contrast, the congruent compounds, in which information about the omission of an outcome was presented in the context of its positive predictor, both biased performance in the opposite direction toward the action earning the outcome whose omission was not predicted and elevated performance above baseline (Figure 2C). Evidence from retardation tests, conducted at the end of the experiment, and the results of a control study assessing the effects of training using only one excitor (i.e., S1S3-O1 and S1S4-O2) confirmed that the effects of S3 and S4 on performance were due to the specific negative predictions established to those stimuli during training and were not due to any direct excitatory learning (see Supplemental Experimental Procedures, Supplemental Discussion, and Figure S2). As such, the current results provide further evidence that the rats can encode both factual and counterfactual action-outcome mappings and are able to flexibly alter their choice of action based on predictive information regarding the occurrence or omission of a specific outcome.

Experiment 3: Forward and Backward Conditioning Exert Opposing Effects on Choice

Finally, we used backward conditioning, in which the outcome is presented before the stimulus to reverse the usual causal relationship, thereby allowing rats to learn that the stimulus predicts the absence of any immediate outcome. The design of this experiment is illustrated in Figure 3A. We used three groups of naive rats. The first group received standard Pavlovian training in a forward manner similar to that used in the previous studies and during which two stimuli, S1 and S2, predicted the grain and sugar outcomes, O1 and O2. For the other two groups of rats, either the outcomes were presented immediately prior to the stimuli with a 0-s delay between them (i.e., simultaneous with stimulus onset), or the outcomes were presented 10 s prior to their respective stimuli to ensure a more distinct backward

A

Groups	Pavlovian Day 1-8	Instrumental Day 9-16	Retraining Day 17	Choice Test Day 18
Forward (For)	S1-O1 S2-O2	A1-O1 A2-O2	S1-O1 S2-O2	S1 A1 vs. A2 S2 A1 vs. A2
Backward 0 (Bck0)	O1-S1 O2-S2	A1-O1 A2-O2	O1-S1 O2-S2	S1 A1 vs. A2 S2 A1 vs. A2
Backward 10 (Bck10)	O1---S1 O2---S2	A1-O1 A2-O2	O1---S1 O2---S2	S1 A1 vs. A2 S2 A1 vs. A2

B

Figure 3. Generating Negative Predictions by Backward Conditioning

(A) The backward conditioning procedure used to generate negative predictors. The outcomes were delivered immediately after the stimuli in group Forward (For, n = 24), immediately before the stimuli in group Backward0 (Bck0, n = 8), and 10 s before the stimuli in group Backward10 (Bck10, n = 16). Abbreviations are the same as those in the legend of Figure 1.

(B) At test, groups Forward and Backward0 exhibited similar behavior (F < 0.1), as the stimuli biased choice toward the Same action, which earned the outcome with which the stimuli had been previously associated ($F_{(1,45)} = 10.97$, p < 0.05). The remaining group Backward10 displayed an opposite pattern of performance ($F_{(1,45)} = 10.97$, p < 0.05). In this group, a stimulus directed choice toward the Different action, which was associated with the absence of the outcome that the stimuli also signaled would not occur ($F_{(1,45)} = 7.78$, p < 0.05). No difference was found between the various groups in the absence of the stimuli (Fs < 2.5). Error bars denote ±1 SEM.

arrangement. All rats then received training on the two lever press actions for the outcomes, such that A1 earned O1 and A2 earned O2 as in the previous studies. Finally, choice between A1 and A2 was again assessed in an extinction test in which no outcomes were delivered. To examine the effect of the stimuli on choice, we periodically presented either S1 or S2 during the test for each of the groups.

We anticipated that both the forward and the simultaneous presentation of outcomes and stimuli would provide sufficient information to bias choice toward the action associated with the predicted outcome; i.e., that S1 would bias choice toward A1 and S2 would bias choice toward A2 on test. The important question was how the rats would respond in the 10-s delay backward-paired group. As both stimuli predict the absence of any subsequent outcome, it is possible they would generally inhibit choice. Alternatively, as the stimuli were presented so as to encourage learning about the absence of one outcome in particular (O1 in the case of S1 and O2 in the case of S2), we hoped that the stimulus-specific predictive learning would reveal evidence that the rats had encoded the counterfactual mappings during instrumental training and that they could use these mappings to inform their decision-making on test; i.e., the presence of S1 would lead the rats to predict the absence of O1 and thus select A2—the action associated with no O1 during training.

The results from the first two training phases and from a direct test of the negative predictions produced by the backward con-

ditioning are presented in Figure S3. The results from the test phase are presented in Figure 3B as described previously (see Supplemental Experimental Procedures). As in previous experiments, the predictive stimuli established using either forward stimulus-outcome pairings or the zero-delay procedure were found to provide sufficient positive predictive information for the stimuli to bias choice toward the action that in training delivered the outcome predicted by the stimulus. In contrast, presenting the outcomes 10 s prior to the stimuli produced the opposite effect on choice. As was observed in the previous studies, in this condition, the stimuli predicting the absence of a specific outcome resulted in the rats choosing the action also associated with the absence of the outcome (Figure 3B). Together with the prior two studies, this finding provides consistent evidence that during the instrumental training, the rats encoded both the factual and the counterfactual action-outcome relationships. They not only encoded the specific outcome delivered by an action, they clearly also encoded the outcome that the action did not produce and were able to use this information flexibly on test to guide choice.

DISCUSSION

Counterfactual reasoning appears to be a ubiquitous capacity in humans that emerges very early in life [18–20]. This ubiquity has been interpreted as evidence that such reasoning is not a specialized function but, rather, plays a more general role in the ongoing regulation of action and particularly goal-directed action [21, 22]. As such, various theories of counterfactual reasoning have emphasized both its dependency on fundamental structural elements of our transactions with the environment [23] and its functional role as a necessary component in the management and coordination of ongoing behavioral strategy [6]. In line with its ubiquity and its necessity for the regulation of adaptive behavior, the current study provides evidence that the ability to encode counterfactual relationships and to use them to guide choice can be extended to rats and hence to animals other than humans. By examining the influence of positive and negative predictions on choice between actions, we observed that stimuli predicting the absence of particular outcomes appeared as able to influence action selection as stimuli predicting their presence. These stimuli were generated by various inhibitory conditioning procedures and were found both to bias choice away from the action earning the inhibited outcome and to elevate the performance of an alternative action that was associated with the absence of that outcome. The reversal of the effect of positive prediction by a negative predictor provides direct evidence that rats do not only encode the factual relationships between actions and outcomes but can also encode counterfactual relationships, i.e., the outcomes that their actions do not produce. Furthermore, the flexibility with which the rats applied these relationships during a single brief test suggests that they not only encode these relationships but can reason counterfactually. When faced with information about the absence of a specific outcome, they did not simply stop choosing but immediately modified their performance to select the alternative action.

The reversal of choice and the increase in the selection of alternative actions produced by negative predictions in the

current experiments constitute a challenge for formal theories of choice and decision-making, particularly currently popular computational theories such as those derived from reinforcement learning [24–27]. Generally, these theories hold in common the view that preference for one action over others is determined by its relative value. This value is typically calculated according to the probability and desirability of the earned outcome and is conferred on the action via the encoding of a specific action value. The probability and desirability of other outcomes, i.e., those that are not delivered by the action, are not considered in this evaluative process. It follows, therefore, that, in the current experiments, these theories predict a general reduction in responding and indifference in choice when a stimulus signaling the absence of a particular outcome is presented. Certainly, no current variant of these theories predicts that the performance of alternative actions will be elevated by these stimuli as observed in the present experiments. This elevation is, however, anticipated by assuming the encoding of inhibitory action-outcome relationships during instrumental conditioning, i.e., by assuming that the rats encode not only the outcomes that actually occur but also those that do not occur as a consequence of an action. Indeed, we propose that, as animals learn that one action leads to a particular outcome, e.g., A1→O1 or A2→O2, they also learn the implied counterfactual relationship—that the action does not lead to outcomes that can be earned by other actions, i.e., A1→O1 not O2 and A2→O2 not O1.

Although the conditions governing inhibitory action-outcome associations are yet to be determined (see Supplemental Discussion), there is a further issue regarding whether the flexible use of such associations qualifies as "reasoning." We have framed our analysis in terms of counterfactual reasoning because these action-outcome mappings affected choice without further learning, but their general adaptive significance also implies such reasoning processes are generally engaged during instrumental conditioning; e.g., besides the current demonstrations, they would appear to be central to instrumental avoidance, for which the ability to rapidly select actions that do not lead to specific consequences has obvious adaptive advantages [28]. Whether rational inference should be considered a process over and above the implementation of excitatory and inhibitory action-outcome associations in choice performance is, however, an open question [29, 30]. In many cases, it can be shown that associative explanations often sidestep rather than implement cognitive explanations of behavior [31], whereas in others, most notably in the case of associative theories of goal-directed action, it has been more compellingly argued that constraining associative principles within a specific processing architecture can manifest rational cognition [31], and perhaps the same can be said about the influence of predictive learning on choice between goal-directed actions in the current studies.

Although evidence of causal learning in animals has been met with considerable skepticism [8], the evidence that rats can make decisions based on predictive information, whether it relates to the presence or absence of significant outcomes or the necessity of making an intervention [10], suggests that they can reflect both on the causal consequences of their actions and also on the consequences that their actions do not cause. Rather than disengage and inhibit actions altogether, rats appear able to actively choose an action that they have learned is not associated with an outcome that currently available information suggests will not occur, something they could only do given an appreciation of, and the ability to act on the basis of, the counterfactual state of affairs.

REFERENCES

1. Sloman, S. (2009). Causal Models: How People Think About the World and Its Alternatives. (Oxford: Oxford University Press).

2. Hume, D. (1978). A Treatise of Human Nature. (Oxford: Oxford University Press).

3. Mackie, J.L. (1974). The Cement of the Universe: A Study of Causation. (Oxford: Oxford University Press).

4. Lewis, D. (1973). Causation. J. Philos. 70, 556–567.

5. Pearl, J. (2000). Causality Models, Reasoning, and Inference | Philosophy of science. (Cambridge: Cambridge University Press).

6. Epstude, K., and Roese, N.J. (2008). The functional theory of counterfactual thinking. Pers. Soc. Psychol. Rev. 12, 168–192.

7. Roese, N.J. (1997). Counterfactual thinking. Psychol. Bull. 121, 133–148.

8. Penn, D.C., and Povinelli, D.J. (2007). Causal cognition in human and nonhuman animals: a comparative, critical review. Annu. Rev. Psychol. 58, 97–118.

9. Taylor, A.H., Miller, R., and Gray, R.D. (2012). New Caledonian crows reason about hidden causal agents. Proc. Natl. Acad. Sci. USA 109, 16389–16391.

10. Blaisdell, A.P., Sawa, K., Leising, K.J., and Waldmann, M.R. (2006). Causal reasoning in rats. Science 311, 1020–1022.

11. Rescorla, R.A. (1994). Control of instrumental performance by Pavlovian and instrumental stimuli. J. Exp. Psychol. Anim. Behav. Process. 20, 44–50.

12. Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. Anim. Learn. Behav. 22, 1–18.

13. Dickinson, A., and Balleine, B.W. (2000). Causal cognition and goal-directed action. In Evolution of Cognition, C. Hayes and L. Huber, eds. (Cambridge: MIT press), pp. 185–204.

14. Holland, P.C. (1984). Differential effects of reinforcement of an inhibitory feature after serial and simultaneous feature negative discrimination training. J. Exp. Psychol. Anim. Behav. Process. 10, 461–475.

15. Rescorla, R.A., and Lolordo, V.M. (1965). Inhibition of avoidance behavior. J. Comp. Physiol. Psychol. 59, 406–412.

16. Rescorla, R.A. (1970). Reduction in the effectiveness of reinforcement after prior excitatory conditioning. Learn. Motiv. 1, 372–381.

17. Delamater, A.R., LoLordo, V.M., and Sosa, W. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. Learn. Behav. *31*, 393–402.

18. Summerville, A., and Roese, N.J. (2008). Dare to compare: fact-based versus simulation-based comparison in daily life. J. Exp. Soc. Psychol. *44*, 664–671.

19. Perner, J., Sprung, M., Zauner, P., and Haider, H. (2003). Want that is understood well before say that, think that, and false belief: a test of de Villiers's linguistic determinism on German-speaking children. Child Dev. *74*, 179–188.

20. Gilovich, T., Wang, R.F., Regan, D., and Nishina, S. (2003). Regrets of action and inaction across cultures. J. Cross Cult. Psychol. *34*, 61–71.

21. Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. Philos. Trans. R. Soc. Lond. B Biol. Sci. *369*, 20130474.

22. Guajardo, N.R., Parker, J., and Turley-Ames, K. (2009). Associations among false belief understanding, counterfactual reasoning, and executive function. Br. J. Dev. Psychol. *27*, 681–702.

23. Pearl, J. (2013). Structural counterfactuals: a brief introduction. Cogn. Sci. *37*, 977–985.

24. Solway, A., and Botvinick, M.M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. Psychol. Rev. *119*, 120–154.

25. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711.

26. Lee, D., Seo, H., and Jung, M.W. (2012). Neural basis of reinforcement learning and decision making. Annu. Rev. Neurosci. *35*, 287–308.

27. Sutton, R.S., and Barto, A.G. (1998). Introduction to Reinforcement Learning, First Edition. (Cambridge: MIT Press).

28. Seligman, M.E., and Johnston, J. (1973). A cognitive theory of avoidance learning. In Contemporary Approaches to Conditioning and Learning, F.J. McGuigan and D.B. Lumsden, eds. (Washington: Winston and Sons), pp. 69–110.

29. Heyes, C. (2012). Simple minds: a qualified defence of associative learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 2695–2703.

30. Mitchell, C.J., De Houwer, J., and Lovibond, P.F. (2009). The propositional nature of human associative learning. Behav. Brain Sci. *32*, 183–198, discussion 198–246.

31. Dickinson, A. (2012). Associative learning and animal cognition. Philos. Trans. R. Soc. Lond. B Biol. Sci. *367*, 2733–2742.