Model-free and model-based learning processes in the updating of explicit and implicit evaluations

Benedek Kurdi[a,1], **Samuel J. Gershman**[a,b], and **Mahzarin R. Banaji**[a,1]

[a]Department of Psychology, Harvard University, Cambridge, MA 02138; and [b]Center for Brain Science, Harvard University, Cambridge, MA 02138

Contributed by Mahzarin R. Banaji, February 9, 2019 (sent for review November 27, 2018; reviewed by Molly J. Crockett and Leor M. Hackel)

Evaluating stimuli along a good–bad dimension is a fundamental computation performed by the human mind. In recent decades, research has documented dissociations and associations between explicit (i.e., self-reported) and implicit (i.e., indirectly measured) forms of evaluations. However, it is unclear whether such dissociations arise from relatively more superficial differences in measurement techniques or from deeper differences in the processes by which explicit and implicit evaluations are acquired and represented. The present project (total *N* = 2,354) relies on the computationally well-specified distinction between model-based and model-free reinforcement learning to investigate the unique and shared aspects of explicit and implicit evaluations. Study 1 used a revaluation procedure to reveal that, whereas explicit evaluations of novel targets are updated via model-free and model-based processes, implicit evaluations depend on the former but are impervious to the latter. Studies 2 and 3 demonstrated the robustness of this effect to (*i*) the number of stimulus exposures in the revaluation phase and (*ii*) the deterministic vs. probabilistic nature of initial reinforcement. These findings provide a framework, going beyond traditional dual-process and single-process accounts, to highlight the context-sensitivity and long-term recalcitrance of implicit evaluations as well as variations in their relationship with their explicit counterparts. These results also suggest avenues for designing theoretically guided interventions to produce change in implicit evaluations.

Implicit Association Test | implicit evaluations | implicit social cognition | model-free vs. model-based learning | reinforcement learning

The human mind continuously assigns subjective value to information encountered in the environment (1). Such evaluations of humans, abstract concepts, and physical objects are crucial to structuring thinking, feeling, and behavior. A wealth of research conducted over the past 30 y has shown that evaluations can be revealed not only via traditional self-report measures (i.e., explicit evaluations) but also via more indirect measures of response interference (i.e., implicit evaluations) (2). One such measure, the Implicit Association Test (IAT) (2), indexes relative evaluations of two targets (e.g., social groups, individuals, or objects) by using a comparison of response latencies across two speeded sorting tasks: a first sorting task in which one of the targets shares a response key with positive items and the other target shares a response key with negative items, and a second sorting task in which the mapping of targets to valences is reversed. For instance, an implicit evaluation is inferred based on the speed and accuracy to associate the concept "flower" (e.g., tulip, daisy) with pleasant attributes (e.g., angel, success) while associating the concept "insect" (e.g., bug, fly) with negative attributes (e.g., devil, failure) vs. the opposite pairing of flower with negative attributes and insect with positive attributes.

Implicit evaluations have been shown to predict behavior beyond their explicit counterparts in a range of consequential settings, including intergroup relations, consumer choice, psychopathology, and close relationships (3, 4). For instance, implicit evaluations of African American targets measured at the level of geographic areas predict police brutality (5), implicit evaluations of products predict product usage and brand recognition (6), implicit evaluations of the self and self-injury predict suicidal behavior (7), and implicit evaluations of one's romantic partner predict long-term relationship success (8). As such, understanding the processes by which implicit evaluations emerge and are updated is not only of theoretical interest across several areas of psychology but also of considerable practical and societal importance.

Implicit and explicit evaluations differ from each other in terms of the method by which they are measured. Implicit evaluations are usually indexed by tasks that bypass effortful control, with typical measures involving speeded responses to preselected pairs of stimuli. By contrast, explicit evaluations are usually measured by using self-report (e.g., via responses to Likert items). Dominant dual-process theories of evaluation posit that, beyond differences in measurement, explicit and implicit evaluations also differ from each other in more profound ways. Crucially, explicit and implicit evaluations are commonly hypothesized to originate from fundamentally different learning processes. Specifically, the learning processes giving rise to explicit evaluations are posited to be flexible and rule-governed and to rely on propositional information, whereas the learning processes giving rise to implicit evaluations are posited to be slow and gradual and to rely on associative regularities encountered in the environment (9–11).

Even though this dual-process perspective on evaluative learning has inspired much empirical work on the acquisition and change of explicit and implicit evaluations, it suffers from some notable shortcomings. First, in opposition to the theory, it has been

Significance

Explicit (i.e., self-reported) and implicit (i.e., indirectly measured) evaluations are central to organizing social cognition and drive behavior in intergroup relations, consumer choice, psychopathology, and close relationships. Here we use a distinction rooted in animal learning and now formalized in computer science to show that explicit evaluations are responsive to model-free and model-based reinforcement learning, whereas implicit evaluations are sensitive to the former but impervious to the latter. This result, which demonstrates a partial dissociation in computations underlying explicit vs. implicit evaluations, accounts for a wide range of existing findings and productively generates predictions for future research. Moreover, it suggests that enduring change in implicit evaluations may be achieved via retraining model-free value representations across a number of different contexts.

Author contributions: B.K., S.J.G., and M.R.B. designed research; B.K. performed research; B.K. analyzed data; and B.K., S.J.G., and M.R.B. wrote the paper.

Reviewers: M.J.C., Yale University; and L.M.H., Stanford University.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Open Science Framework (https://osf.io/f8pg3/).

[1]To whom correspondence may be addressed. Email: kurdi@g.harvard.edu or mahzarin_banaji@harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820238116/-/DCSupplemental.

Published online March 12, 2019.

repeatedly demonstrated that implicit evaluations can be flexibly updated via purely verbal instructions and that such updating need not involve direct experience with any stimulus (12). The preponderance of such findings has prompted some to abandon a dual-process perspective on evaluative learning altogether and to replace it with a model of evaluative learning that relies on a single propositional process (13–15).

Second, dual-process theories of evaluation, as currently conceived, are difficult to falsify, and the same applies to single-process alternatives. For instance, whether learning is quick or slow is a matter of judgment, and, as such, researchers with different theoretical commitments may make widely divergent inferences from the very same data. Moreover, it is unclear what kind of empirical evidence would be sufficient to discern whether evaluative information is represented in the form of conceptual associations (e.g., flower–good), as posited by dual-process theories of evaluation, or equivalence relationships (e.g., "flowers are good"), as posited by propositional alternatives.

Third, implicit evaluations exhibit a host of characteristics that are not accounted for by theories that claim that they are subserved by a set of enduring associative representations accumulated over time. For instance, implicit evaluations have been shown to be situationally malleable (16). Specifically, implicit evaluations respond to motivational states, such as nicotine deprivation, thirst, and hunger (17), as well as to higher-order goals, such as the goal to be egalitarian (18). At the same time, contrary to the prediction by a single-process propositional perspective, implicit evaluations are not indiscriminately sensitive to verbal interventions that have been demonstrated to shift explicit evaluations (19).

Finally, dual-process and single-process theories of evaluation both make extreme predictions about the relationship that should emerge between explicit and implicit evaluations. According to dual-process theories, explicit and implicit evaluations are subserved by different learning processes, and, as such, any convergence between the two is unexpected. By contrast, according to single-process theories, explicit and implicit evaluations are subserved by a single learning process, and, as such, there is no reason to expect them to diverge. However, the overwhelming majority of empirical data fall between the two extremes: explicit and implicit evaluations are typically correlated with each other but are rarely redundant (20), with the magnitude of the correlation modulated by the domain. For instance, explicit and implicit evaluations of political candidates have been found to be highly correlated, whereas explicit and implicit evaluations of racial groups often show considerably lower levels of correlation (20).

Even though current dual-process theories of evaluation do not explain all of the available evidence on the updating of implicit evaluations, they are not easily falsifiable, and are silent on a host of phenomena related to the malleability of implicit evaluations, it may be premature to abandon the class of dual-process theories altogether (21). In this paper, we use reinforcement learning algorithms, which originate from the study of animal learning (22) and now play an important role in computer science and artificial intelligence (23), to provide a test of the manner in which explicit and implicit evaluations are acquired and updated. Specifically, we investigate whether explicit and implicit evaluations are equally or differentially sensitive to model-free and model-based learning. If both respond in similar ways, we can conclude that, despite differences in measurement techniques, the representations underlying explicit and implicit evaluations are likely similar to each other. If, on the contrary, the two differ in their sensitivity to model-free and model-based learning, we can conclude that the data are more suggestive of differences in learning and representation.

In a reinforcement learning framework, an agent interacts with its environment and, via such interaction, pursues two distinct but interrelated goals: (*i*) to learn about the actions that produce the largest amount of long-term reward and (*ii*) to adjust behavior in line with this learning. Given the generality of this framework, rewards can range from primary reinforcers such as food or sex to more abstract rewards like points in a game or even social reinforcers like smiles and group inclusion. To solve the reinforcement learning problem just described (i.e., to maximize long-term reward), an agent must create internal representations of the subjective value associated with taking different actions. Most important for the present purposes, such representations can be created in two fundamentally different ways (23): by using model-free or model-based algorithms. The distinction between model-free and model-based processes has already been used with great success to elucidate phenomena across diverse areas of psychology, including moral cognition (24), impression formation (25), and addiction (26). Here we use it to study the acquisition and shift of implicit evaluations.

Even though model-free and model-based algorithms solve the same problem of maximizing long-term reward, they differ from each other in the way they learn and the kind of information they are able to represent. In the present studies (Fig. 1), participants made choices between two fictitious social targets (Laapians vs. Niffians). Depending on their choice, they were then exposed to an intervening stimulus (a horizontal vs. a vertical bar), followed by a win or a loss. In this setting, the goal of both model-free and model-based algorithms is to learn whether, in the long run, choosing a Laapian target or choosing a Niffian target is the more advantageous action. However, they accomplish this goal in fundamentally different ways.

Model-free algorithms operate over an unordered list of actions, each of which is associated with a positive or negative scalar value. For instance, in the present studies, the model-free system may represent two actions ("choose Laapian" vs. "choose Niffian") and, in the absence of prior learning, associate an initial value of zero with each. Over the course of the task, learning unfolds incrementally and based on direct experience: each time the agent performs an action (e.g., choosing a Laapian target), it updates the value associated with that action based on its outcomes. For instance, if choosing a Laapian target results in a positive outcome (e.g., winning points), the agent increases the value associated with that action, and if it results in a negative outcome (e.g., losing points), the agent decreases the value associated with it. Incremental updating is performed until the prediction error is reduced to zero, i.e., there is no more discrepancy between the reward expected and actually received. Such an algorithm is computationally cheap: it creates action–value pairs, which constitute a highly compressed representation of the history of rewards. However, the simplicity of this algorithm comes at a cost of reduced flexibility. Specifically, action–value pairs can be updated only upon performing an action. Moreover, given that specific outcomes of actions (e.g., Laapians leading to horizontal bars leading to wins) are not represented, the model-free system has no way to modulate its behavior based on higher-level goals (e.g., "I want to get to the horizontal bar").

Unlike model-free algorithms that operate exclusively over action–value pairs (e.g., Laapian, +5; Niffian, −5), model-based algorithms operate over a considerably richer cognitive map of the environment that represents the specific outcomes of actions. For instance, in the context of the present experiments, a model-based agent would create a causal model linking first-stage stimuli to second-stage stimuli [e.g., "whenever I choose Laapians, I get to a horizontal bar" or $P(\text{horizontal} \mid \text{Laapian}) = 1$] and second-stage stimuli to rewards [e.g., "whenever I see a horizontal bar, I win five points" or $P(+5 \mid \text{horizontal}) = 1$]. This representation involves considerably more detail than the highly compressed representation created by model-free learning. Crucially, by virtue of representing specific outcomes of actions [e.g., $P(\text{horizontal} \mid \text{Laapian}) = 1$], model-based learning can bypass the trial-by-trial updating that characterizes model-free

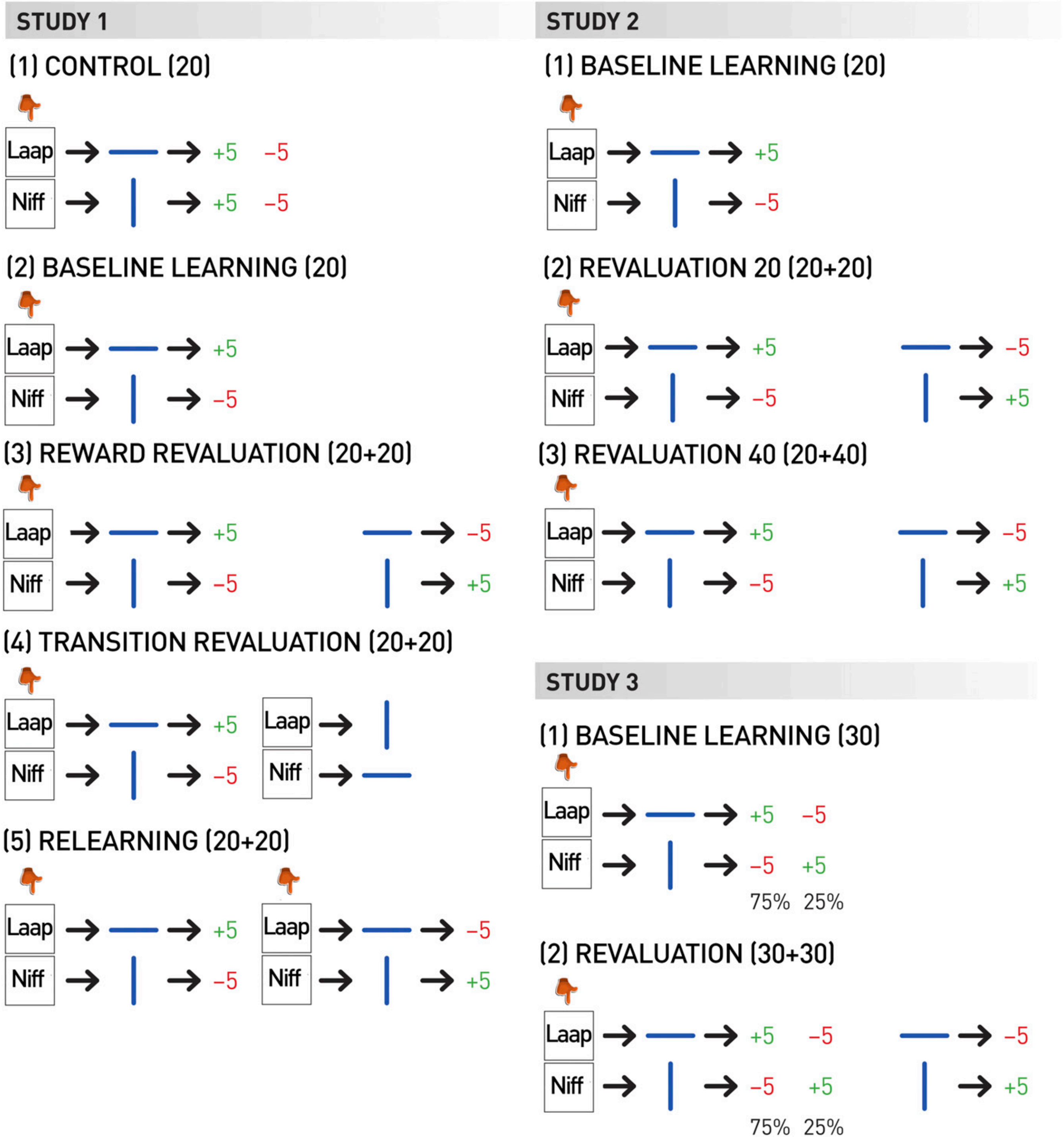

Fig. 1. Overview of the learning-phase procedure for studies 1–3. The number of trials (for each part of the learning phase where applicable) is noted after the name of the condition. A hand symbol indicates a choice made by the participant. The assignment of first-stage stimuli [Laapians (Laap) vs. Niffians (Niff)] to second-stage stimuli (horizontal vs. vertical bars) as well as the assignment of second-stage stimuli to outcomes (+5 vs. −5 points) was counterbalanced across participants. Transitions between first-stage stimuli, second-stage stimuli, and outcomes were deterministic, with the exception of the control condition in study 1, in which second-stage stimuli were randomly followed by wins or losses, and study 3, in which initial learning in both conditions was probabilistic (one second-stage stimulus followed by wins on 75% of trials and by losses on 25% of trials and the other second-stage stimulus followed by losses on 75% of trials and by wins on 25% of trials). In the control (study 1) and baseline learning (studies 1–3) conditions, all dependent measures (transition memory, explicit evaluation, and implicit evaluation) were administered following the learning phase. In all remaining conditions, transition memory and explicit evaluation items were administered following each part of the learning phase, whereas implicit evaluations were measured only after the second part of the learning phase.

learning. Instead, model-based representations enable mental simulation of different courses of action by considering the goal to be achieved (e.g., getting to the horizontal bar) and the probabilities with which different actions (e.g., choosing Laapians) can bring about the desired goal. As such, unlike model-free algorithms, model-based algorithms are highly flexible. However, their flexibility comes at a cost: effortful planning over different courses of action may be prohibitively complex and time-consuming, especially if the number of potential actions and outcomes to plan over is large.

Nonhuman animals (27, 28) and humans (29, 30) have been shown to exhibit model-free and model-based learning. Under some conditions, model-free and model-based algorithms are difficult to tease apart because they converge on the same behavioral output. For instance, learning in the baseline learning condition of the present experiments (Fig. 1) can be accomplished via model-free or model-based processes: that is, participants may simply learn to associate higher value with the stimulus that led to a positive outcome in the past (e.g., Laapian, +5; Niffian, −5). Alternatively, participants may explicitly represent the structure of the task, i.e., create a mental model of transition probabilities [e.g., $P(\text{horizontal} \mid \text{Laapian}) = 1$, $P(+5 \mid \text{horizontal}) = 1$, $P(\text{vertical} \mid \text{Niffian}) = 1$, $P(-5 \mid \text{vertical}) = 1$].

However, the results of model-free and model-based learning can diverge when the environment changes in such a way as to modify the motivational relevance of a known stimulus. Specifically, a paradigm commonly referred to as reward revaluation has been used to discern whether nonhuman animals (27, 28) and humans (29, 30) rely on model-free or model-based learning. In these studies, participants undergo initial learning that establishes that an action (e.g., pressing a lever or choosing an abstract image) is rewarding (e.g., results in the participant receiving food pellets or monetary rewards). In a second stage, the rewarding quality of the reinforcer is eliminated in the absence of the participant taking any relevant action: for instance, the previously satiating food pellets are paired with illness or the previously winning image is paired with monetary loss.

The signature difference between model-free and model-based learning is then revealed when participants are again allowed to take the action that produced the previously rewarding outcome. Model-free algorithms are backward-looking and inflexible and, as such, can update values associated with an action only after that action has been performed and a reward has been experienced. Therefore, participants pursuing a purely model-free strategy will continue to consistently perform the action (e.g., pressing the lever or selecting the image) even following revaluation because of its history of producing reward. By contrast, model-based algorithms are forward-looking and flexible and, as such, have the ability to incorporate information about the new state of affairs. As such, participants pursuing a purely model-based strategy now expect that ingesting the food pellets will induce sickness or choosing the abstract image will result in a loss, and will thus consistently refrain from performing the previously reinforced action. In such paradigms, human participants usually pursue a mixture of both strategies (29, 30); however, importantly, any decrease in the tendency to perform the initially reinforced action the can be interpreted as reflecting the contributions of model-based learning.

The Present Project

The present project has three interrelated goals. First, a large body of research has provided evidence that value representations in humans, as revealed by explicit measures of self-report, can be updated on the basis of the rewards received as a result of interacting with a given stimulus (29–33). However, despite a similarly large body of research investigating the effects of mere exposure (34), Pavlovian learning (35), approach–avoidance training (36), and verbal instructions (12) on implicit evaluations, to our knowledge, the effects of reinforcement learning, i.e., rewarding or punishing participants for taking actions involving motivationally relevant stimuli, on implicit evaluations has never been investigated. As such, the first goal of the present project is to establish whether implicit evaluations of novel stimuli can be effectively shifted via this form of learning.

Second, and most important, the present project was designed to probe whether explicit and implicit evaluations of novel targets are equally sensitive to model-free and model-based learning. As mentioned here earlier, explicit evaluations revealed by self-report have been demonstrated to be responsive to both model-free and model-based learning (29–33). However, whether their implicit counterparts are characterized by the same or different patterns of updating is an open question, with different theoretical perspectives and lines of empirical work making opposing predictions about the pattern of data that should emerge.

A prediction of convergence between explicit and implicit evaluations can be made based on propositional theories of implicit evaluation (13–15), widely replicated patterns of empirical data, and the nature of the IAT (2), which was used as the dependent measure in all three studies reported in the present paper. Specifically, propositional theories of implicit evaluation posit that explicit and implicit evaluations do not differ in underlying learning processes or mental representations. As such, if the present revaluation paradigm successfully shifts explicit evaluations of a stimulus, such learning should also be reflected by implicit measures of evaluation. Second, in studies involving novel targets, such as the present one, patterns of convergence between explicit and implicit evaluations are common because participants do not have access to any information about the targets other than the information provided by the experimenter (12). Moreover, in this setting, pressures to act in a socially desirable manner, known to result in dissociations between explicit and implicit evaluations (20), are unlikely to operate. Finally, the IAT, unlike most implicit measures, requires participants to hold two to four categories in working memory while completing the task. This feature of the procedure may activate model-based representations.

On the contrary, dual-process theories of evaluative learning (9–11) generally predict dissociations between explicit and implicit evaluations. Specifically, reward revaluation, which can be accomplished via model-based but not via model-free processes, should be expected to shift explicit but not implicit evaluations: the use of information that is not represented in precompiled form to evaluate a stimulus may require effortful processing characteristic of the explicit system. Moreover, recent empirical work has revealed that placing participants under cognitive load while performing a reinforcement learning task shifts them toward reliance on the computationally cheap model-free system and away from reliance on the computationally expensive model-based system (31). Most implicit measures of evaluation, such as the IAT, place participants under similar cognitive constraints given that they involve responding under time pressure. As such, this difference in the availability of cognitive resources across the explicit vs. implicit evaluation tasks may also contribute to a pattern of explicit–implicit dissociation.

Third, to the extent that the present project provides evidence for a dissociation between explicit and implicit evaluations, it is important to demonstrate that such dissociation is accounted for by a computational difference between model-free and model-based reinforcement learning. However, in a standard revaluation paradigm, use of model-free vs. model-based strategies is confounded with primacy vs. recency. Specifically, model-free learning would be revealed via reliance on initially learned information and model-based learning would be revealed via successful updating. Importantly, it has been shown that implicit evaluations, including implicit evaluations of novel targets, may be difficult to change once they are in place (37). As such, any convincing claim about explicit–implicit dissociation being a result of the model-free vs. model-based distinction has to involve a condition controlling for the temporal confound inherent in the reward revaluation paradigm.

Study 1

Design. The experiment in study 1 consisted of a learning phase and one or two test phases (depending on condition). In the learning phase, participants interacted with two novel groups (Laapians vs. Niffians) and received rewards (positive points) or punishments (negative points) as a result of their choice behavior. In the test phases, participants provided forced-choice judgments probing (*i*) transition memory and (*ii*) the value of the Laapian and Niffian targets (explicit evaluation), followed by (*iii*) an IAT (2) probing implicit evaluation of the same targets.

Crucially, for the learning phase of the experiment, participants (final n = 1,740) were assigned to one of five between-subjects conditions (Fig. 1). In the control and baseline learning conditions, the learning phase consisted of a single part, whereas in the reward revaluation, transition revaluation, and relearning conditions, the learning phase consisted of two parts.

Across all five learning conditions, the first part of the learning phase required participants to complete 20 learning trials on which they made a choice between a Laapian and a Niffian target (first-stage stimuli; *Materials and Methods*). Depending on their choice, participants were exposed to a horizontal or a vertical bar (second-stage stimulus), followed by a positive outcome (+5 points) or a negative outcome (−5 points). Participants were instructed to maximize the points received. The relationship between first-stage and second-stage stimuli was deterministic in all five conditions (e.g., Laapians were always followed by horizontal bars and Niffians by vertical bars). In the control condition, second-stage stimuli were randomly followed by wins or losses, thus providing a measure of relative preference at baseline. In all four remaining conditions, the transition between second-stage stimuli and rewards was deterministic (e.g., horizontal bars were always followed by wins and vertical bars by losses).

In the reward revaluation, transition revaluation, and relearning conditions, the first part of the learning phase was followed by a second part, also consisting of 20 trials. In the reward revaluation condition, the transition between second-stage stimuli and rewards was reversed compared with the first part of the learning phase (without participants making any choices or experiencing any first-stage stimuli). In the transition revaluation condition, the transition between first-stage and second-stage stimuli was reversed compared with the learning phase (without participants making any choices or experiencing any rewards). The relearning condition was similar to the reward revaluation condition in that the transition between second-stage stimuli and rewards was reversed; however, unlike in the reward revaluation condition, participants experienced the full transition structure from first-stage stimuli to second-stage stimuli to rewards and, rather than passively observing stimuli, they made choices between Laapian and Niffian targets.

In the control and baseline learning conditions, the learning phase was followed by (*i*) a set of explicit transition memory items probing memory for the transition between first-stage and second-stage stimuli, (*ii*) a set of explicit evaluation items probing self-reported subjective value assigned to each target (Laapians vs. Niffians), and (*iii*) an IAT probing relative implicit evaluation of Laapians vs. Niffians. In the reward revaluation, transition revaluation, and relearning conditions, the explicit transition memory and explicit evaluation items were administered twice, once after the first part of the learning phase and once after the second part of the learning phase. However, to prevent participant fatigue, the IAT was administered only once, following the second part of the learning phase.

The logic of the statistical analyses reported here later is as follows. First, a comparison involving the control and baseline learning conditions can be used to establish whether the rewards and punishments used in the present task were effective in shifting participants' explicit and implicit evaluations of first-stage stimuli. Explicit evaluations have been demonstrated to shift as a result of similar manipulations numerous times (29–33); the present project also provides a test of whether a binary choice between two targets, followed by rewards and punishments, can successfully shift implicit evaluations as measured by the IAT.

Second, a crucial comparison involving the baseline learning and reward revaluation conditions can be used to probe whether explicit and implicit evaluations are sensitive to model-based learning. As noted here earlier, successful updating of subjective value in the reward revaluation condition is commonly interpreted to rely only on model-based processes given that the second part of the learning phase did not involve any experience with first-stage stimuli. Similar to the first comparison, the effectiveness of reward revaluation in shifting explicit evaluations has already been demonstrated (29–33); by contrast, to our knowledge, whether reward revaluation can shift implicit evaluations has not been investigated before.

Third, a comparison involving the baseline learning and transition revaluation conditions can be used to probe whether explicit and implicit evaluations are sensitive to a different kind of change in the environment. The predictions for this comparison are less straightforward than for the reward revaluation condition given that updating in this condition may occur via model-free or model-based processes or a combination of both: model-based updating may be performed if participants use their explicit model of the task to cognitively link the second-stage stimuli to rewards (as experienced in the first part of the learning phase). However, because second-stage stimuli were paired with wins and losses in the first part of the learning phase, they might act as valenced stimuli themselves, thus enabling model-free learning (akin to second-order conditioning).

Fourth, a comparison involving the baseline learning and relearning conditions can be used to help disambiguate the results of the reward revaluation condition by revealing whether implicit evaluations are differentially sensitive to (*i*) model-free vs. model based learning or (*ii*) initial learning vs. subsequent updating (i.e., a primacy effect) (37). Specifically, if implicit evaluations were to be insensitive to model-based learning, such insensitivity would be reflected by statistically equivalent responding in the baseline learning and reward revaluation conditions (as detailed earlier). However, this pattern of responding may also be the result of implicit evaluations being generally more responsive to initial learning than to updating based on novel information. If this is the case, and implicit evaluations are generally impervious to updating, no difference would be expected between the baseline learning and relearning conditions given that the relearning condition, just like the reward revaluation condition, involves initial learning followed by updating. By contrast, if the defining difference is between model-free and model-based processes, the relearning condition, unlike the reward revaluation condition, should show change given that, in the former, unlike in the latter, learning can be accomplished via model-free processes.

Results. The pattern of results obtained with explicit evaluation as the dependent measure (Fig. 2) was in line with expectations formulated on the basis of similar studies conducted in the past (29–33) and, as such, underscores the soundness of the design and manipulations.

Specifically, baseline learning was found to be effective in shifting explicit evaluations compared with the control condition [$t(548.86) = 9.88$, $P < 0.0001$, Bayes Factor in favor of the alternative hypothesis (BF_{10}) = 3.40×10^{18}, Cohen's $d = 0.82$], thus establishing the general effectiveness of the learning task used in the present study (for further evidence see *SI Appendix*, *Supplementary Studies S1* and *S2*). Also in line with expectations, reward revaluation was effective in shifting explicit evaluations compared with the baseline learning condition [$t(474.09) = 14.49$, $P < 0.0001$, $BF_{10} = 5.89 \times 10^{38}$, Cohen's $d = 1.22$], thus replicating the widely observed finding that explicit evaluations respond to

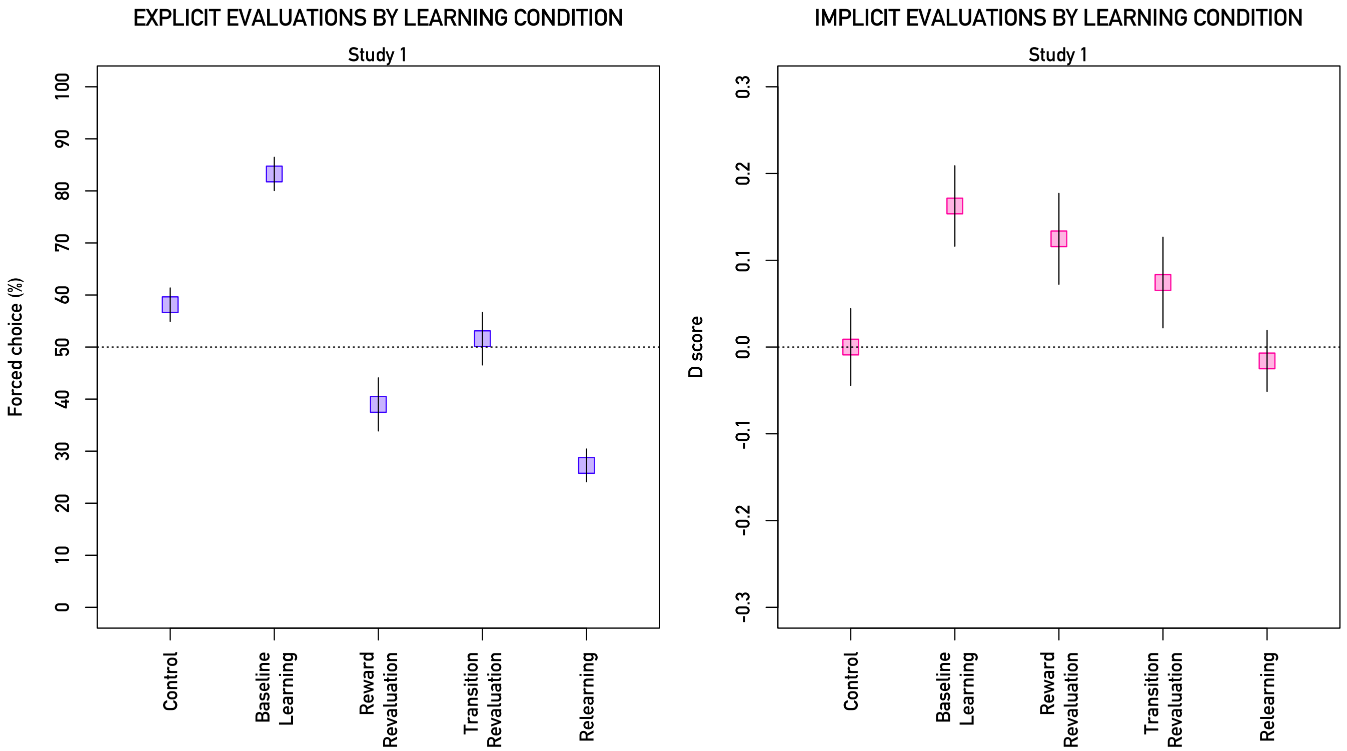

Fig. 2. Study 1 (n = 1,740): mean explicit and implicit evaluations by learning condition. For explicit evaluations (*Left*), the y axis shows percentage of responses in line with initial learning; for implicit evaluations (*Right*), the y axis shows IAT D scores (44) computed such that higher values indicate responses in line with initial learning. For explicit and implicit evaluations, effects of revaluation or relearning are revealed by values closer to 0% or negative D scores, respectively. In the control condition, responses indicating preference in favor of Laapians over Niffians were arbitrarily coded as positive. For visualization purposes, IAT scores have been mean-centered by using the mean of the control condition. Error bars show 95% CIs.

model-based learning (for further evidence see *SI Appendix, Supplementary Study S1*).* A similar result was observed for the comparison involving baseline learning and transition revaluation [$t(502.54) = 10.44$, $P < 0.0001$, $\mathrm{BF}_{10} = 5.06 \times 10^{20}$, Cohen's d = 0.86], which should not be surprising given that such updating could have occurred via model-free or model-based processes. Finally, although not of major theoretical relevance for the present purposes, the relearning condition was also found to effectively shift explicit evaluations compared with the baseline learning condition [$t(793.91) = 24.55$, $P < 0.0001$, $\mathrm{BF}_{10} = 1.62 \times 10^{85}$, Cohen's d = 1.59; for further evidence see *SI Appendix, Supplementary Study S3*].

Given that learning in the baseline learning and relearning conditions could have been accomplished in a model-free or a model-based way, we investigated (*i*) whether explicit evaluations differed from neutrality at chance level responding to the transition memory item (a signature of model-free processes) and (*ii*) whether accuracy of transition memory predicted explicit evaluations (a signature of model-based processes) in each condition. As revealed by a significant intercept, explicit evaluations differed from neutrality at chance level responding to the transition memory item in the baseline learning condition [b = 0.90, $t(307) = 11.76$, $P < 0.0001$] and in the relearning condition [b = 0.47, $t(573) = 6.07$, $P < 0.0001$], thus providing evidence for the operation of model-free processes. At the same time, accurate transition memory positively predicted explicit evaluations in both conditions [$b = 0.43$, $t(307) = 8.77$, $P < 0.0001$; and $b = 0.38$, $t(573) = 8.83$, $P < 0.0001$], revealing the contribution of model-based processes to the acquisition of explicit evaluations.

A comparison involving the control and baseline learning conditions revealed that implicit evaluations, like explicit evaluations, were sensitive to reinforcement learning: scores on the IAT exhibited significant change away from control as a result of the rewards received in the baseline learning condition [$t(565.06) = 4.35$, $P < 0.0001$, $\mathrm{BF}_{10} = 9.11 \times 10^{2}$, Cohen's $d = 0.36$; for further evidence see *SI Appendix, Supplementary Studies S1* and *S2*). The crucial comparison in this experiment involved the baseline learning vs. reward revaluation conditions given that this comparison establishes whether implicit evaluations responded to model-based reinforcement learning. This comparison provided evidence in favor of the null hypothesis [$t(569.05) = 1.06$, P = 0.287, Bayes Factor in favor of the null hypothesis (BF_{01}) = 6.22, Cohen's d = 0.09], suggesting that implicit evaluations are impervious to model-based updating (for further evidence see *SI Appendix, Supplementary Study S1*). In line with the expectation that updating in the transition revaluation condition may emerge from model-free or model-based processes, we found weak evidence that the transition revaluation condition may have differed from the baseline learning condition [$t(591.29) = 2.47$, $P = 0.013$, $\mathrm{BF}_{10} = 1.85$, Cohen's $d = 0.20$].

Finally, given that we found no updating compared with baseline in the reward revaluation condition, a comparison involving the baseline learning and relearning conditions can be used to establish whether such lack of updating occurred as a result of a general primacy effect or the more specific effect of implicit evaluations being impervious to model-based, but not

*In this condition, as well as the transition revaluation and relearning conditions, participants completed two sets of explicit evaluation items, one following the first part of the learning phase (initial learning) and one following the second part of the learning phase (revaluation or relearning). To ensure compatibility of analyses across conditions and across explicit and implicit measures, here were report comparisons between the baseline learning condition and the second set of explicit evaluation items completed following revaluation or relearning. In the *SI Appendix*, we report additional within-participant analyses comparing the first set vs. the second set of explicit evaluation items in the reward revaluation, transition revaluation, and relearning conditions. These within-participant analyses reinforce the conclusions reported here.

model-free, updating. The baseline learning and relearning conditions were found to significantly differ from each other [$t(649.58) = 6.04$, $P < 0.0001$, $BF_{10} = 2.40 \times 10^6$, Cohen's $d = 0.42$], suggesting that already-established implicit evaluations can be effectively updated provided that the updating can be performed via model-free mechanisms. As such, this result eliminates a general primacy effect as an explanation for the present findings (for further evidence see *SI Appendix, Supplementary Study S3*).

Studies 2 and 3

Design. Study 1 has provided initial evidence that, unlike their explicit counterparts, implicit evaluations are impervious to model-based learning. Studies 2 and 3 were designed to provide a direct replication of this result as well as to produce evidence about its generality (Fig. 1).

In line with our primary focus on the sensitivity of implicit evaluations to model-based learning, studies 2 and 3 consisted only of baseline learning and reward revaluation conditions. In study 2 (final $n = 245$), in addition to the baseline learning condition, two versions of the reward revaluation condition were implemented. In the first version (revaluation 20 condition), participants were exposed to 20 revaluation trials. As such, this condition provides a direct replication of the reward revaluation condition from study 1. In the second version (revaluation 40 condition), participants were exposed to 40, rather than 20, revaluation trials. As such, in this condition, participants experienced twice as many trials in the second part of the learning phase (i.e., revaluation) than in the first part (i.e., initial learning). If a comparison of implicit evaluations across the baseline learning and revaluation 40 conditions reveals no difference, this would suggest that the lack of updating observed in study 1 was likely a result of the insensitivity of implicit evaluations to model-based learning rather than a lack of sufficient training in the second part of the learning phase.

In study 3 (final $n = 369$), reinforcement in the baseline learning condition and in the first part of the revaluation condition (i.e., initial learning) was probabilistic rather than deterministic, with one of the targets (e.g., Laapians) followed by wins 75% of the time and the other target (e.g., Niffians) followed by losses 75% of the time. To provide a conservative test of the null hypothesis of model-based learning being ineffective in shifting implicit evaluations, revaluation was deterministic. This study was designed to probe whether the insensitivity of implicit evaluations to model-based learning, as established by study 1, may be modulated by the ambiguity of the initially received evaluative information. As such, the reinforcement contingencies in this study were more ecologically realistic than contingencies in studies 1 and 2, in which one of the targets was deterministically followed by wins and the other target was deterministically followed by losses.

Results. In study 2, explicit evaluations shifted significantly as a result of revaluation in the revaluation 20 condition [$t(124.17) = 8.31$, $P < 0.0001$, $BF_{10} = 2.04 \times 10^{12}$, Cohen's $d = 1.33$] and the revaluation 40 condition [$t(115.47) = 7.42$, $P < 0.0001$, $BF_{10} = 1.44 \times 10^{10}$, Cohen's $d = 1.22$; *SI Appendix*, Fig. S1], thus replicating the results obtained in study 1. In study 3, explicit evaluations were also found to shift significantly as a result of reward revaluation, although the evidence in favor of change was considerably weaker than in studies 1 and 2 [$t(351.96) = 2.16$, $P = 0.031$, $BF_{10} = 1.10$, Cohen's $d = 0.23$; *SI Appendix*, Fig. S2].

Crucially, replicating the results of study 1, implicit evaluations were found to be impervious to reward revaluation in study 2 (*SI Appendix*, Fig. S1) and study 3 (*SI Appendix*, Fig. S2). Specifically, study 2 provided evidence in favor of the null hypothesis when the number of trials was the same across the first and second parts of the learning phase [baseline learning vs. revaluation 20 conditions, $t(154.74) = -0.87$, $P = 0.381$, $BF_{01} = 4.19$, Cohen's $d = -0.14$]. A similar result was obtained in the revaluation 40 condition in which the number of revaluation trials was double the number of the initial learning trials [$t(152.71) = -0.51$, $P = 0.612$, $BF_{01} = 5.28$, Cohen's $d = -0.08$]. Implicit evaluations also remained impervious to reward revaluation in study 3, demonstrating that their insensitivity to model-based learning does not depend on the deterministic vs. probabilistic nature of initial reinforcement [$t(366.99) = 0.32$, $P = 0.747$, $BF_{01} = 8.26$, Cohen's $d = -0.03$].

Results Combined Across Experiments

Bayesian meta-analyses were conducted to obtain an aggregate measure of differences across the baseline learning and reward revaluation conditions in studies 1–3. Explicit evaluations were found to be sensitive to model-based learning ($BF_{10} = 2.58 \times 10^{43}$, Cohen's $d = 0.87$, 95% highest-density interval (HDI) = [0.77; 0.99]; *SI Appendix*, Fig. S3), replicating previous work (29–33). By contrast, implicit evaluations were found to be impervious to model-based learning ($BF_{01} = 13.18$, Cohen's $d = 0.03$, 95% = [−0.08; 0.15]; *SI Appendix*, Fig. S4). Additional meta-analyses conducted only with participants who had perfect transition memory revealed the same pattern of results (*SI Appendix*, Figs. S5 and S6), suggesting that lack of updating in implicit evaluations did not result from an erroneous representation of the structure of the environment.

To further compare the relative importance of transition memory in shaping implicit vs. explicit evaluations, we conducted a small-sample–corrected robust metaregression (38) with correlation between transition memory and evaluation as the dependent measure, a fixed effect for type of evaluation (implicit vs. explicit), and a random effect for study and condition to account for dependency in the data. For implicit measures, no relationship was found between transition memory and evaluation [$b = 0.013$, 95% CI (−0.028 to 0.054), $t(7.91) = 0.75$, $P = 0.476$]. By contrast, for explicit measures, transition memory positively and significantly predicted evaluations [$b = 0.237$, 95% CI (0.072–0.390), $t(7.91) = 3.29$, $P = 0.011$]. As such, this meta-analysis provides additional correlational evidence for the idea that explicit, but not implicit, evaluations are responsive to model-based learning.

Discussion

We conducted three experiments relying on the distinction between model-free and model-based reinforcement learning (23, 29) and involving novel stimuli to arrive at a better understanding of the updating of implicit (indirectly measured) evaluations. Model-free algorithms are backward-looking, incremental, and computationally cheap: they adjust the value of an action upon experiencing its motivationally relevant outcomes. By contrast, model-based algorithms are forward-looking, flexible, and computationally expensive: they perform planning over a causal model of the environment to choose the best course of action in light of current goals.

The model-free vs. model-based distinction seemed ideal as a theoretical basis for this investigation because, unlike existing dual-process and single-process theories of evaluation, it is computationally well-specified: the signatures of model-free vs. model-based processes can be revealed in a so-called revaluation paradigm (27–30). In this paradigm, subjective evaluation of a well-known and previously rewarding stimulus is measured after the stimulus loses its rewarding quality. Change in choice behavior as a result of this new information reveals the operation of model-based learning, whereas persistence of the old choice behavior is characteristic of a model-free strategy. Explicit evaluations have been known to reflect a combination of model-free and model-based processes (29–33), and here we replicate this result.

The contribution of the present project is twofold. First, we show that, in addition to other forms of evaluative learning such as mere exposure (34), Pavlovian learning (35), approach–avoidance training (36), and verbal instructions (12), implicit evaluations of stimuli, similar to their explicit counterparts, are amenable to updating as a result of reinforcement learning, i.e., experience with the positive and negative outcomes of actions involving those stimuli. Second, we demonstrate a commonality and a difference in the computations underpinning the updating of explicit vs. implicit evaluations via reinforcement learning: just like explicit evaluations, implicit evaluations were found to be responsive to model-free processes at baseline and following initial model-free learning with different reinforcement contingencies. However, unlike their explicit counterparts, implicit evaluations were insensitive to model-based learning.

Such dissociation between explicit and implicit evaluations is surprising for a number of reasons. First, our own previous work has shown that implicit evaluations can shift in the face of propositional processes traditionally thought of as uniquely influencing explicit evaluations, suggesting underlying commonality in learning (12). However, here we provide clear evidence for a theoretically meaningful explicit–implicit dissociation. Second, to conduct a conservative test of dissociation, explicit items were always administered to participants before the IAT, and, as such, responding on the IAT could have been influenced not only by the learning manipulations but also by responding on the explicit measure of evaluation. However, no spillover effects were observed, suggesting separate underlying representations. Third, the IAT, used as a measure of implicit evaluation in the present studies, involves explicit categorization of stimuli, which may have been predicted to activate model-based value representations, but the data showed no such evidence. Fourth, all three experiments involved novel social groups as targets. This feature should have minimized social desirability concerns, which are known to contribute to explicit–implicit dissociations in tests of real social targets (20).[†]

However, although the present studies created a pattern of dissociation between explicit and implicit evaluations, it should be noted that a reinforcement learning perspective, unlike a traditional dual-process perspective (9–11), does not make an unqualified prediction of explicit–implicit dissociations, for multiple reasons. First, in many situations, including the baseline learning condition of the present studies, model-free and model-based algorithms converge on the same value representation. Second, as demonstrated by the present studies, explicit and implicit evaluations can both be updated by model-free processes. This shared learning process should generally lead to some degree of association between explicit and implicit evaluations. Third, recent research has shown that model-free and model-based learning need not be antagonistic: on the contrary, a model of the environment can be used to modulate model-free value representations via simulated experience (33, 39). Future work may test this idea in the context of implicit evaluations by imposing a delay between the revaluation and test phases of the experiment.

In addition, the distinction between model-free and model-based learning processes provides a theoretical framework to explain why certain interventions, even those that do not involve valenced feedback upon performing an action, can successfully shift implicit evaluations, whereas others seem to be ineffective. Among 17 interventions implemented in a recent large-scale collaboration, with individual investigators submitting their chosen intervention, eight shifted implicit evaluations of African American subjects toward neutrality, whereas nine produced no change (19). Among the eight interventions that were effective, five were clearly better characterized as model-free: they included direct experience with African American exemplars paired with positive stimuli or outcomes [e.g., evaluative conditioning (35)]. The remaining three manipulations that were effective required a mental model given that they were based on verbal instructions rather than direct experience; however, this mental model was of the simplest possible form: $P(\text{positive} \mid \text{African American}) = 1$ and $P(\text{negative} \mid \text{white American}) = 1$. This group of interventions included a vivid story in which the protagonist was assaulted by a white American and saved by a black American, as well as two manipulations involving implementation intentions (40). From a reinforcement learning perspective, it could be argued that the causal model involved in the latter group of interventions is sufficiently simple to be able to train model-free values almost immediately (see also ref. 12).

By contrast, among the nine ineffective interventions, eight involved a complex causal model of the environment, including a model of another person's mind, a model of a positive encounter with an outgroup member, and a model of racial injustice. Crucially, unlike the successful interventions described here earlier, this set of interventions did not provide participants with precompiled value representations (e.g., African American, good; white American, bad) that could be activated quickly and effortlessly while responding under time pressure on an implicit task. Given this time pressure, participants may not have had sufficient opportunity to discern what modulation of existing value representations a complex causal model would imply. Finally, the only model-free intervention that remained ineffective involved pairings of both black and white Americans with (*i*) positive and negative facial expressions and (*ii*) positive and negative feedback. Given the nature of reinforcement provided in this intervention, no change in the model-free values associated with each target should be expected.

This perspective on the results reported in ref. 19 is consistent with the idea that evaluative representations acquired in ways other than via reinforcement learning may generally be able to effectively drive responding on implicit measures such as the IAT only if they are sufficiently compressed to enable automatic activation under time pressure. Future work will be able to offer more systematic tests of this idea. For instance, the model-free vs. model-based distinction underpinning the present project may be used to probe whether implicit evaluations are amenable to revaluation in a Pavlovian setting (41). Moreover, when it comes to purely language-based learning (12, 13), the present results suggest that the effectiveness of verbal statements in updating implicit evaluations may be moderated by the complexity of the propositional reasoning required to assign the appropriate truth value to those verbal statements or, in the terminology of reinforcement learning, by the complexity of the implied causal model.

Moreover, the present results, as well as a general reinforcement learning framework, provide a perspective on what is usually described as the sensitivity of implicit evaluations to higher-order goals (17, 18). In studies of this kind, activation of a goal (e.g., hunger, achievement, or egalitarianism) leads to a modulation of implicit evaluations such that objects that can contribute to achieving the goal are temporarily evaluated more positively until the goal is successfully completed. These findings are seemingly at odds with the present results, given that, as mentioned earlier, only model-based, and not-model free, value representations can be modulated in the face of higher-order goals.

However, the contradiction between both perspectives may be illusory. One group of variables investigated in this set of studies (including nicotine deprivation, thirst, and hunger) are more appropriately described as motivational states rather than goals. Sensitivity of model-free reinforcement learning to motivational states is compatible with the theoretical formulation of model-free algorithms (42) and empirical findings (22): based on past

[†] Even though the stimuli were introduced to participants as "names [...] com[ing] from two groups," it should be pointed out that the stimuli were not imbued with any further social meaning, nor did we replicate the present experiments using stimuli explicitly described as nonsocial. As such, we leave a direct comparison of social vs. nonsocial and novel social vs. real social stimuli in the present paradigm to future work.

experience, a model-free learner can represent multiple value estimates associated with the same action (e.g., smoking a cigarette). Smoking a cigarette in a nicotine-deprived state is highly rewarding, and smoking a cigarette in a nicotine-satiated state is much less so. Accordingly, over time, a smoker should learn to associate higher value with cigarettes in the former compared with the latter context and activate the appropriate value representation depending on their current motivational state.

A second set of variables used in this literature can be described as genuine higher-order goals (such as achievement or egalitarianism). However, the general finding involving such goals is that they modulate responding on implicit measures only to the extent that participants have protracted past experience with them (e.g., professional athletes or chronic egalitarians). In a reinforcement learning framework, such past experience is equivalent to having accumulated corresponding model-free value representations over time. These model-free representations can then be activated automatically upon encountering the relevant motivational state without such activation requiring genuine goal-directed behavior involving effortful planning over a causal model. As such, in line with our observation described earlier, motivational states and goals seem to modulate implicit evaluations only to the extent that they provide a precompiled value representation that can be activated automatically and effortlessly during an implicit task.

The present findings demonstrating the sensitivity of implicit evaluations to model-free learning and their insensitivity to model-based learning may be expanded upon in a number of ways in future work. For instance, as mentioned earlier, model-free learning is inherently state-dependent, which may provide an explanation for the highly contextualized nature of implicit evaluation (16) as well as its resistance to long-term change (43). By mapping out the space of relevant states and providing model-free training across a large number of them, change in implicit evaluations may become more robust, enduring, and generalizable. Moreover, as mentioned earlier, the present studies were designed to produce a dissociation between model-free and model-based learning; however, recent work on offline updating of model-free value representations via model-based algorithms (33, 39) suggests that model-based interventions may also be successfully used to shift implicit evaluations. Beyond these specific proposals for future work, it is our hope that the theoretical framework outlined here will generally inspire much insightful theorizing and empirical research on when, how, and why implicit evaluations change.

Materials and Methods

Institutional Approval and Informed Consent. All studies reported here were granted ethical approval by the Committee on the Use of Human Subjects at Harvard University. Participants provided informed consent at the beginning of each study.

Participants. Participants in all studies were American adult volunteers recruited from the Project Implicit educational Web site (implicit.harvard.edu). Exclusion criteria are reported in the *SI Appendix*.

Learning Phase. In the initial part of the learning phase of studies 1 and 2, participants completed 20 forced choice trials. In study 3, the number of forced choice trials was increased to 30. On each trial, a Laapian stimulus (randomly selected from *Caalap*, *Feelslap*, *Gabeelap*, *Ineelap*, and *Maasolap*) and a Niffian stimulus (randomly selected from *Ibbonif*, *Jabbunif*, *Lebbunif*, *Mettanif*, and *Oballnif*) were presented side-by-side on the screen. Participants selected the left-hand stimulus by pressing the E key or the right-hand stimulus by pressing the I key. The side on which Laapian and Niffian stimuli were presented was randomly selected for each trial. Following participants' choice, a second-stage stimulus (horizontal or vertical bar) was displayed. When participants had pressed the space bar, the second-stage stimulus was removed and a reward (+5 or −5) appeared. The next trial started upon pressing the space bar. The transition from first-stage stimuli (Laapians vs. Niffians) to second-stage stimuli (horizontal vs. vertical bars) to rewards (+5 vs. −5 points) was counterbalanced across participants.

The learning phase of the control (study 1) and baseline learning (studies 1–3) conditions consisted of only the initial learning described earlier. In the reward revaluation (studies 1–3), transition revaluation (study 1), and relearning (study 1) conditions, a second part followed. In study 1, the second part of the learning phase consisted of 20 trials; in study 2, it consisted of 20 or 40 trials (depending on condition); and, in study 3, it consisted of 30 trials. In the reward revaluation conditions, participants were exposed to a randomly selected second-stage stimulus (horizontal or vertical bar) on each trial. When they had pressed the C key, a reward (+5 or −5 points) was revealed. The next trial started upon the participant pressing the space bar. The transition revaluation condition was similar, with the exception that participants were exposed to first-stage stimuli (Laapians or Niffians) and, upon pressing the C key, a second-stage stimulus (horizontal or vertical bar) appeared. The second part of the learning phase in the relearning condition of study 1 was identical to the first part, with the exception that the transition from second-stage stimuli (horizontal vs. vertical bars) to rewards (+5 vs. −5 points) was reversed.

Explicit Evaluation. Explicit evaluation items were identical to the forced-choice trials used in the first part of the learning phase; however, on these trials, participants received no feedback. Participants in the control and baseline learning conditions completed a single set of four explicit evaluation items, whereas participants in the reward revaluation, transition revaluation, and relearning conditions completed two sets of four explicit evaluation items: one set following the first part of the learning phase (initial learning) and a second set following the second part of the learning phase (revaluation or relearning). Responses on each set of explicit evaluation items were summed (1 = normatively accurate response, 0 = normatively inaccurate response) to create an index of explicit evaluation.

Transition Memory. Transition memory items were identical to the explicit evaluation items, with the exception that participants were asked to select the first-stage stimulus (Laapian vs. Niffian) that would lead to a certain second-stage stimulus (horizontal vs. vertical bar) rather than to a positive outcome. The administration of transition memory items followed the same schedule as the administration of explicit evaluation items. Responses on each set of transition memory items were summed (1 = normatively accurate response, 0 = normatively inaccurate response) to create an index of transition memory.

Implicit Evaluation. Implicit evaluations were measured by using a standard five-block IAT (2). The categories were "Laapians" and "Niffians" and the attributes were "good" and "bad." Category items were identical to the items used during the learning phase. Good attribute items included *love*, *peace*, *joy*, *happy*, *peace*, *glory*, and *lucky*; bad attribute items included *hate*, *war*, *devil*, *bomb*, *bitter*, *agony*, and *grief*. Implicit evaluations were calculated by using the improved scoring algorithm (44) such that higher *D*-scores indicate evaluations in line with initial learning. Further details of the IAT procedure are reported in the *SI Appendix*.

Statistical Analyses. All statistical analyses were conducted in the R statistical computing environment. The R code for all analyses, as well as data files (including trial-level IAT data), are freely available from the Open Science Framework (https://osf.io/f8pg3/) (45). The Bayesian *t* tests and Bayesian meta-analyses were performed by using the BayesFactor package (46). The small-sample–corrected robust variance metaregression was conducted by using the robumeta package (47).

ACKNOWLEDGMENTS. We thank Harrison Satcher for assistance with study coding and Kornél Sóti for assistance with preparing Fig. 1. This work was supported by the Dean's Competitive Fund for Promising Scholarship of the Harvard University Faculty of Arts and Sciences (to M.R.B.) and an Undergraduate Research Scholar award from the Institute of Quantitative Social Science at Harvard University (to M.R.B.).

1. Allport GW (1935) Attitudes. *A Handbook of Social Psychology*, ed Murchison C (Clark Univ Press, Worcester, MA), pp 798–844.
2. Greenwald AG, McGhee DE, Schwartz JLK (1998) Measuring individual differences in implicit cognition: The Implicit Association Test. *J Pers Soc Psychol* 74: 1464–1480.
3. Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR (2009) Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J Pers Soc Psychol* 97:17–41.
4. Kurdi B, et al. (December 13, 2018) Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *Am Psychol*, 10.1037/amp0000364.

5. Hehman E, Flake JK, Calanchini J (2017) Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Soc Psychol Personal Sci* 9: 393–401.
6. Maison D, Greenwald AG, Bruin RH (2004) Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *J Consum Psychol* 14: 405–415.
7. Nock MK, Banaji MR (2007) Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *J Consult Clin Psychol* 75:707–715.
8. McNulty JK, Olson MA, Meltzer AL, Shaffer MJ (2013) Though they may be unaware, newlyweds implicitly know whether their marriage will be satisfying. *Science* 342: 1119–1120.
9. Smith ER, DeCoster J (2000) Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Pers Soc Psychol Rev* 4:108–131.
10. Rydell RJ, McConnell AR (2006) Understanding implicit and explicit attitude change: A systems of reasoning analysis. *J Pers Soc Psychol* 91:995–1008.
11. Strack F, Deutsch R (2004) Reflective and impulsive determinants of social behavior. *Pers Soc Psychol Rev* 8:220–247.
12. Kurdi B, Banaji MR (2017) Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes? *J Exp Psychol Gen* 146:194–213.
13. De Houwer J (2014) A propositional model of implicit evaluation. *Soc Personal Psychol Compass* 8:342–353.
14. De Houwer J, Hughes S (2016) Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Soc Cogn* 34:480–494.
15. Mitchell CJ, De Houwer J, Lovibond PF (2009) The propositional nature of human associative learning. *Behav Brain Sci* 32:183–198, discussion 198–246.
16. Blair IV (2002) The malleability of automatic stereotypes and prejudice. *Pers Soc Psychol Rev* 6:242–261.
17. Ferguson MJ, Bargh JA (2008) Evaluative readiness: The motivational nature of automatic evaluation. *Handbook of Approach and Avoidance Motivation*, ed Elliot AJ (Psychology Press, New York), pp 287–304.
18. Moskowitz GB (2014) The implicit volition model: The unconscious nature of goal pursuit. *Dual-Process Theories of the Social Mind*, eds Sherman JW, Gawronski B, Trope Y (Guilford, New York), pp 400–422.
19. Lai CK, et al. (2014) Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *J Exp Psychol Gen* 143:1765–1785.
20. Nosek BA (2005) Moderators of the relationship between implicit and explicit evaluation. *J Exp Psychol Gen* 134:565–584.
21. Evans JSBT, Stanovich KE (2013) Dual-process theories of higher cognition: Advancing the debate. *Perspect Psychol Sci* 8:223–241.
22. Dickinson A, Balleine B (2002) The role of learning in the operation of motivational systems. *Stevens' Handbook of Experimental Psychology: Learning, Motivation and Emotion*, eds Stevens SS, Pashler HE (Wiley, New York), pp 497–534.
23. Sutton RS, Barto AG (1998) *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, MA).
24. Crockett MJ (2013) Models of morality. *Trends Cogn Sci* 17:363–366.
25. Hackel LM, Doll BB, Amodio DM (2015) Instrumental learning of traits versus rewards: Dissociable neural correlates and effects on choice. *Nat Neurosci* 18:1233–1235.
26. Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nat Neurosci* 8:1481–1489.
27. Adams CD, Dickinson A (1981) Instrumental responding following reinforcer devaluation. *Q J Exp Psychol B* 33:109–121.
28. Dickinson A (1985) Actions and habits: The development of behavioural autonomy. *Philos Trans R Soc B* 308:67–78.
29. Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711.
30. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215.
31. Otto AR, Gershman SJ, Markman AB, Daw ND (2013) The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychol Sci* 24:751–761.
32. Fu W-T, Anderson JR (2008) Dual learning processes in interactive skill acquisition. *J Exp Psychol Appl* 14:179–191.
33. Gershman SJ, Markman AB, Otto AR (2014) Retrospective revaluation in sequential decision making: A tale of two systems. *J Exp Psychol Gen* 143:182–194.
34. Van Dessel P, Mertens G, Smith CT, De Houwer J (2019) Mere exposureeffects on implicit stimulus evaluation: The moderating role of evaluation task, number of stimulus presentations, and memory for presentation frequency. *Pers Soc Psychol Bull* 45:447–460.
35. Hofmann W, De Houwer J, Perugini M, Baeyens F, Crombez G (2010) Evaluative conditioning in humans: A meta-analysis. *Psychol Bull* 136:390–421.
36. Van Dessel P, De Houwer J, Gast A (2016) Approach–avoidance training effects are moderated by awareness of stimulus–action contingencies. *Pers Soc Psychol Bull* 42: 81–93.
37. Gregg AP, Seibt B, Banaji MR (2006) Easier done than undone: Asymmetry in the malleability of implicit preferences. *J Pers Soc Psychol* 90:1–20.
38. Tipton E (2015) Small sample adjustments for robust variance estimation with meta-regression. *Psychol Methods* 20:375–393.
39. Gershman SJ, Zhou J, Kommers C (2017) Imaginative reinforcement learning: Computational principles and neural mechanisms. *J Cogn Neurosci* 29:2103–2113.
40. Gollwitzer PM (1993) Goal achievement: The role of intentions. *Eur Rev Soc Psychol* 4: 141–185.
41. Dayan P, Berridge KC (2014) Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cogn Affect Behav Neurosci* 14:473–492.
42. Niv Y, Joel D, Dayan P (2006) A normative perspective on motivation. *Trends Cogn Sci* 10:375–381.
43. Lai CK, et al. (2016) Reducing implicit racial preferences: II. Intervention effectiveness across time. *J Exp Psychol Gen* 145:1001–1016.
44. Greenwald AG, Nosek BA, Banaji MR (2003) Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *J Pers Soc Psychol* 85:197–216.
45. Kurdi B, Gershman S, Banaji M (2019) Model-free and model-based learning processes in the updating of explicit and implicit evaluations. Open Science Framework. Available at https://osf.io/f8pg3/. Deposited November 24, 2018.
46. Morey RD, Rouder JN, Jamil T (2015) Package "BayesFactor." Version 0.9.12-4.2. Available at https://cran.r-project.org/web/packages/BayesFactor/BayesFactor.pdf. Accessed December 28, 2018.
47. Fisher Z, Tipton E (2015) robumeta: An R-Package for Robust Variance Estimation in Meta-Analysis. Available at https://arxiv.org/abs/1503.02220v1. Accessed December 28, 2018.