

A computational reinforcement learning account of social media engagement

Björn Lindström^{1,2}, Martin Bellander³, Allen Chang⁴, Philippe N. Tobler⁵, David M. Amodio^{1,6}

1. Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
2. Section for Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm Sweden
3. Center for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
4. Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts, USA
5. Zurich Center for Neuroeconomics, Department of Economics, University of Zürich, Switzerland
6. Department of Psychology and Neural Science, New York University, New York, NY, USA

Corresponding author:

Björn R. Lindström
University of Amsterdam
Department of Psychology
Postbus 15900
1001 NK Amsterdam
The Netherlands
bjorn.r.lindstrom@gmail.com

Abstract

Social media has become the modern arena for human life, with billions of daily users worldwide. The intense popularity of social media is often attributed to a psychological need for social rewards (“likes”), which turns the online world into a “Skinner Box” for the modern human. Yet despite such common portrayals, empirical evidence for social media engagement as reward-based behavior remains scant. We applied a computational approach to directly test whether reward learning mechanisms contribute to social media behavior. We analyzed over one million posts from over 4,000 individuals on several social media platforms, using computational models based on reward reinforcement learning theory. Our results consistently show that human behavior on social media qualitatively and quantitatively conforms to the principles of reward learning. Results further reveal meaningful individual differences in social reward learning on social media, explained in part by variability in users’ tendency for social comparison. Together, these findings support the social reinforcement learning view of social media engagement and offer key new insights into this emergent mode of modern human behavior on an unprecedented scale.

A computational reinforcement learning account of social media engagement

What drives people to engage, sometimes obsessively, with others on social media? In 2018, more than three billion people spent(1), on average, several hours a day on platforms such as Instagram, Facebook, Twitter, and other more specialized forums. This pattern of social media engagement is often described as an addiction, in which people are driven to pursue positive online social feedback(2, 3) to the detriment of direct social interaction and even basic needs like eating and drinking(4, 5).

Although a variety of motives might lead people to use social media(6), the popular portrayal of social media engagement as a “Skinner Box” for the modern human suggests it represents a form of reward reinforcement learning (RL)(7). Yet despite this common portrayal, empirical evidence for social media engagement as reward-based behavior has been elusive. In the present research, we developed and applied a computational approach to large scale online datasets of social media use to directly test if, and how, reward learning mechanisms contribute to social media behavior. In doing so, we sought to provide novel insights into this emergent mode of modern human behavior while testing a learning theory model of real-life human social behavior on an unprecedented scale.

In online social media platforms, feedback on one’s behavior often comes in the form of a “like”—a signal of approval from another user regarding one’s post(2)—which is assumed to function as a social reward. Several lines of research indeed support the idea that “likes” engage similar motivational mechanisms as other, more basic, types of rewards. In humans, brain imaging studies have consistently shown that likes(8, 9), and other social rewards, are processed by neural and computational mechanisms closely overlapping with those processing non-social rewards(10–13). Although neuroscientific studies are largely constrained to the laboratory, such findings suggest that social media use might reflect a type

of reward maximization, similar to what is observed across species in response to non-social rewards. More directly, the receipt of likes on social media have behavioral consequences consistent with reward learning. For example, the number of likes received for a post predicts satisfaction with that post, and in turn, more self-reported happiness(14, 15). Similarly, a user's social media activity increases after a post, suggestive of reward anticipation(16). In addition to its direct effect on reward, the subjective value of likes is also influenced by social comparison in a similar way as non-social rewards(3, 17, 18), suggesting that social rewards, just like non-social rewards(19), are relative, rather than absolute in nature. Together, these studies offer suggestive evidence that social media engagement resembles reward reinforcement.

However, as most studies of online social rewards to date utilize self-report methods(20, 21), direct evidence on whether social reward learning processes can explain behavior on social media is lacking. In addition, results from the few studies that have taken a quantitative approach are mixed. In one study, negative evaluation of a post, a type of social punishment, led to deterioration in the quality of future posts, rather than the improvement predicted by learning theory(22). Yet, in another study, receiving more replies for a post on a specific social media discussion forum predicted a subsequent increase in the time spent on that forum relative to others, in-keeping with learning theory(23). Thus, it remains unclear whether basic mechanisms of reward learning can help explain behavior on social media.

Here, we address the critical question of whether social media engagement can be formally characterized as a form of reward learning. By analyzing more than one million posts from over 4000 individual users on several distinct social media platforms (see Methods), we assessed, using computational modeling, how the putative social rewards received for posts in the past (e.g., the likes received when posting a “selfie”) can help explain future behavior. Our computational modeling approach allowed us to explicitly test

how cross-species reward learning mechanisms contribute to a uniquely human mode of social behavior(24).

Computational learning theory posits specific behavioral patterns that would characterize online behavior as an expression of reward learning. A seminal empirical insight is that when animals (e.g., rodents in a Skinner box) can select the timing of their instrumental responses (e.g., when and how often to press a lever), the latency of responding (the inverse of the response rate) is negatively related to the rate of accrued rewards(25). That is, response latencies are longer when the reward rate is higher. Reinforcement learning theory provides both a normative explanation and a mechanistic machinery for this regularity: the more reward one receives, the shorter the average latency between responses should be, because acting more slowly results in a longer delay to the next reward, and the cost of this delay—the opportunity cost of time—is directly related to the average reward rate(26). As consequence, when animals have learned, through interaction with the environment, that the average reward rate is higher, actions should be made faster because more rewards would be foregone by slower and fewer responses.

Although this RL theory was developed to explain animal behavior in laboratory tasks, on timescales of seconds and minutes, the theoretical relationship between the average reward rate and response latency is not tied to any specific timescale. Consequently, if social media taps into basic learning mechanisms, social media behavior should exhibit the same relationship between response latency—the time between successive social media posts—and the (social) reward rate. In other words, we hypothesized that a type of real-life behavior, on timescales rarely or never investigated in the laboratory, would exhibit a key signature of reward learning. Furthermore, given the intrinsically social nature of social media, and the strong human motive to calibrate one's reward experience relative to the successes of others,

we anticipated that the value of online rewards would be influenced by the social context; that is, people's tendency to compare their own outcomes with those of others.

Results

We quantitatively tested our hypothesis that online social behavior in the form of posts follows principles of reward learning theory in four independent datasets (see Methods) (total $N_{\text{Obs}} = 1,046,857$, $N_{\text{Users}} = 4,168$) with computational modeling. These datasets come from four distinct social media platforms, where people post pictures and, in response, receive social reward in the forms of “likes”. In Study 1 ($N_{\text{Users}} = 2,039$), we test our hypothesis in a large dataset of *Instagram* posts(27). Instagram exemplifies modern social media, with over 800 million registered users, and its format—focused primarily on simple postings and the receipt of likes as feedback—makes it a unique case study. However, because there are significant economic motives on Instagram and similar social media(28), as assessment of reward learning on Instagram may be limited to a degree by fraudulent accounts, “fake likes,” and other strategic uses(29). We therefore replicated and extended Study 1 in Study 2 ($N_{\text{Users}} = 2,127$) with data from three different topic-focused social media sites (discussion forums focused on Men's fashion, Women's fashion, and Gardening, see Methods), where economic motives are less likely. In Studies 1 and 2, respectively, the average number of posts per individual was 418 and 91, across average time spans of 731 and 909 days.

Social rewards predict the rate and latency of social media posting. We conceptualized posting on a social media platform (e.g., Instagram) as free-operant behavior in a Skinner box with one response option (e.g., a single lever), where responses are followed by reward (i.e., likes). As outlined, a key prediction from learning theory for such situations where the agent can decide when to respond is that the latency between responses should be affected by the

average rate of rewards(25, 26). Before formally testing our computational hypothesis, we first evaluated, in two complementary and model-independent ways, whether social media behavior was sensitive to social rewards.

First, we drew inspiration from the classical work in animal learning theory, which established that response rates, an aggregate measure of response latency, follow a saturating positive function (i.e., hyperbolic) of reward rates(25). This relationship, known as the quantitative law of effect(25), is a signature of reward driven behavior. To directly test if social media behavior exhibits this signature, we compared how well a hyperbolic function explained the relationship between likes and response rates relative to a linear function (see the Supplementary Information [SI]). We found that the hyperbolic “quantitative law of effect” explained behavior better than a linear relationship in all four datasets (mean R^2 : *Study 1* = 0.43, *Study 2*: = 0.37, see SI), demonstrating that an aggregate measure of response latencies on social media exhibits a classic signature of reward learning(25).

Second, we defined a high resolution measure of response latency (τ_{Post}) as the time between two successive social media posts (similar to the interval between responses in human laboratory tasks and in animal free-operant behavior, see Figure 1), and tested if τ_{Post} was predicted by the history of likes using Granger causality analysis (see Methods). Granger causality is established if a variable (e.g., likes) improves on the prediction of a second variable (e.g., τ_{Post}) over and above earlier (lagged) values of the second variable in itself. To ascertain the selectivity of this method, we first applied it to simulated data from generative models where the ground truth was known (causality or no causality). We then fine-tuned the analysis parameters (the lag number, see SI) to reliably detect Granger causality in data simulated from our reward learning model, which we introduce next, but not from models without learning (where likes are unrelated to behavior, see SI). Applying this optimized analysis method to the empirical data showed that likes Granger caused τ_{Post} in all four

datasets (*Study 1*: $\tilde{Z} = -23.65$, $p < .0001$; *Study 2*: Men's Fashion: $\tilde{Z} = 3.94$, $p < .0001$; Women's Fashion: $\tilde{Z} = 14.16$, $p < .0001$; Gardening: $\tilde{Z} = 6.78$, $p < .0001$). Together, these results demonstrate that the history of social rewards (i.e., likes) influenced both the rate and the time distribution of social media posting. Such reward sensitivity is a minimal criterion for more formally testing the explanatory power of learning theory.

Modeling the dynamics of online behavior as social reward learning. Having established that social media behavior is sensitive to reward, we next developed a generative model, based on RL theory of free-operant behavior in non-human animals(26). The key principle of this theory is that agents should balance the effort costs of responding and the opportunity costs of passivity to maximize the average net (i.e., gains minus losses) reward rate(26). The consequence is that average response latencies should be shorter when the average reward rate is higher. This prediction holds both when the amount of reward is a direct function of the number of responses (i.e., ratio schedules of reinforcement) and when rewards become available at specific time points (i.e., interval schedules of reinforcement).

Building directly on these principles, our $\bar{R}L$ model specifies how agents adjust the latency of their responses to maximize the average net reward rate \bar{R} (see Methods and SI). Hence, the model provides a formal account of our hypothesis that behavior on social media conforms to basic principles of reward learning. Formally, the model conceptualizes social media use as a sequence of decisions regarding the latency between successive posts, τ_{Post} (Figure 1A)(26), where the agent maximizes the reward rate by adaptively adjusting τ_{Post} after observing each accrued reward. Psychologically, τ_{Post} can be thought of as the accumulation of motivation towards a threshold for posting (in similarity to the boundary in evidence accumulation models of decision-making(30)). The model policy, or threshold, which determines τ_{Post} , is dynamically adjusted based on the net reward prediction error, δ , the difference between the experienced reward and the reference level. The reference level is

determined both by the individual’s *effort cost* sensitivity (e.g., the subjective cost of taking pictures and uploading) and the subjective estimate of the average net reward \bar{R} (26, 31), which determines the *opportunity cost* of slow responding (Figure 1 B-C). Both the effort cost and opportunity cost depend on the response latency, τ_{Post} . In other words, the optimal response latency balances these two costs to maximize the net reward δ (Figure 1 D). The subjective estimate of \bar{R} is updated using the same reward prediction error, thereby reflecting the integration of prediction errors across time(26). In total, the model has three free parameters: learning rate, a ; initial policy, P ; and effort cost sensitivity, C (see Methods).

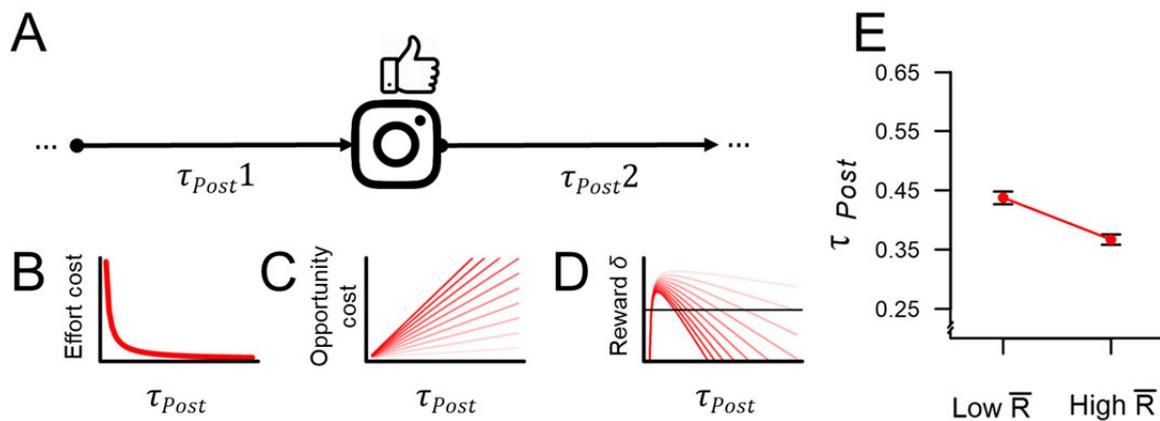


Figure 1. Schematic illustration of the computational hypothesis. (A) The $\bar{R}L$ model describes how τ_{Post} , the latency to next social media post (denoted by the “camera” icon), is shaped by social rewards. Each post is followed by social reward (denoted by the “thumbs up” symbol), which varies in number. The model adjusts the response policy, or threshold, which determines τ_{Post} , to maximize the average net rate of reward. (B) The $\bar{R}L$ model posits an effort cost to responding (e.g., taking pictures, uploading), which decreases as a function of τ_{Post} . The effort cost term penalizes posting in quick succession, because high effort reduces the average reward rate. (C) The opportunity cost of time increases as a function of the average reward rate \bar{R} . The gradient of red lines indicates increasing values of \bar{R} (darker colors represent higher values), and thereby higher opportunity cost. (D) The optimal value of τ_{Post} , which maximizes the net reward δ , varies as function of \bar{R} (darker colors represent higher values). The δ is used to update average reward rate \bar{R} . Note that the optimum, indicated by the peak of the function, moves to shorter response latencies when \bar{R} is higher because the opportunity cost of time increases with \bar{R} . The horizontal line denotes 0. The figure assumes a constant effort cost C . (E) **Simulated model predictions.** The $\bar{R}L$ model predicts that τ_{Post} , the latency between successive social media posts, will be shorter with high compared to low average reward rate, \bar{R} . The simulation involved 1000 synthetic individuals. Errors bars denote 99% CI.

Model predictions. We simulated the model (~250000 data points from 1000 simulated users, with random parameter values, see SI for details) to generate predictions. According to learning theory, τ_{Post} should be lower when the average reward rate is relatively higher. To verify this prediction in a simple manner, we rank-transformed and standardized \bar{R} for each synthetic user and then dichotomized the variable at 0 to produce a qualitative “Low vs High \bar{R} ” predictor (nearly identical results are observed with other definitions, see SI). To facilitate subsequent comparison with empirical analyses, we summarized the simulated data using mixed effects models. These analyses revealed a clear effect of low vs. high \bar{R} on τ_{Post} ($\beta = 0.18$, $SE = 0.007$, $t = 31$, $p < .0001$), as expected. In other words, the model predicts (given the set of simulation parameters) that average response latencies should be ~18% longer when the average reward rate is low versus high (see Figure 1 E). Our empirical analysis of the four social media platforms applied these model-based predictions.

Computational modeling of reward learning in social media behavior

To comprehensively test our hypothesis that online behavior on social media follows principles of basic reward learning, we used computational modeling, statistical analyses, and generative model simulations. We optimized the parameters of the $\bar{R}L$ model for each individual user and quantitatively compared the explanatory value of the $\bar{R}L$ model to a null model without reward learning (see Methods; model estimation and comparison procedure recovered the models with high probability, Figure S1). The null model assumes that posting on social media reflects a stable behavioral tendency (i.e., average response latency, one free parameter), which is not affected by reward. The model comparison provides a direct, quantitative test of reward learning as an explanation for social media use.

Study 1

We first modeled online behavior in the Instagram dataset of Study 1(27). Model comparison showed that the $\bar{R}L$ model accounted better for the time distribution of responses (τ_{Post}) than the model without learning for $\sim 70\%$ of the users (mean individual-level Akaike Information Criterion weight (AIC_w) = .7, 99% CI [0.68, 0.81], one-sample t-test relative to equal AIC_w for the two models: $t(2038) = 23.1$, $p < .0001$). The AIC_w expresses the relative likelihood of one model over another(32). Equivalently, Bayesian random effects model comparison(33) showed that the $\bar{R}L$ model was more common than the model without learning (exceedance probability [xp] = 1), and classified $\sim 70\%$ of individuals as better explained by the $\bar{R}L$ model.

This conclusion was robust to the removal of individuals with especially short or long (e.g., outside the 10th and 90th deciles) average τ_{Post} , or with few (or many) posts (see SI), which confirms that the fit of the $\bar{R}L$ model was not driven by outliers. Similarly, splitting the dataset into four equally sized partitions showed that the $\bar{R}L$ model was the most common in all four partitions (mean AIC_w : 0.68 - 0.73, t-test against equal AIC_w : $t(508) = 9.63-13.9$, $ps < .0001$), which indicates that our conclusion is robust to sample idiosyncrasies and dataset size. Interestingly, we find that the $\bar{R}L$ model fits relatively worse for individuals with many followers, consistent with the possibility that they were not primarily motivated by social rewards (see SI). Moreover, we find that learning models without effort cost (C) or net reward rate (\bar{R}) parameters provide worse accounts of the data (SI and Table S1 for details).

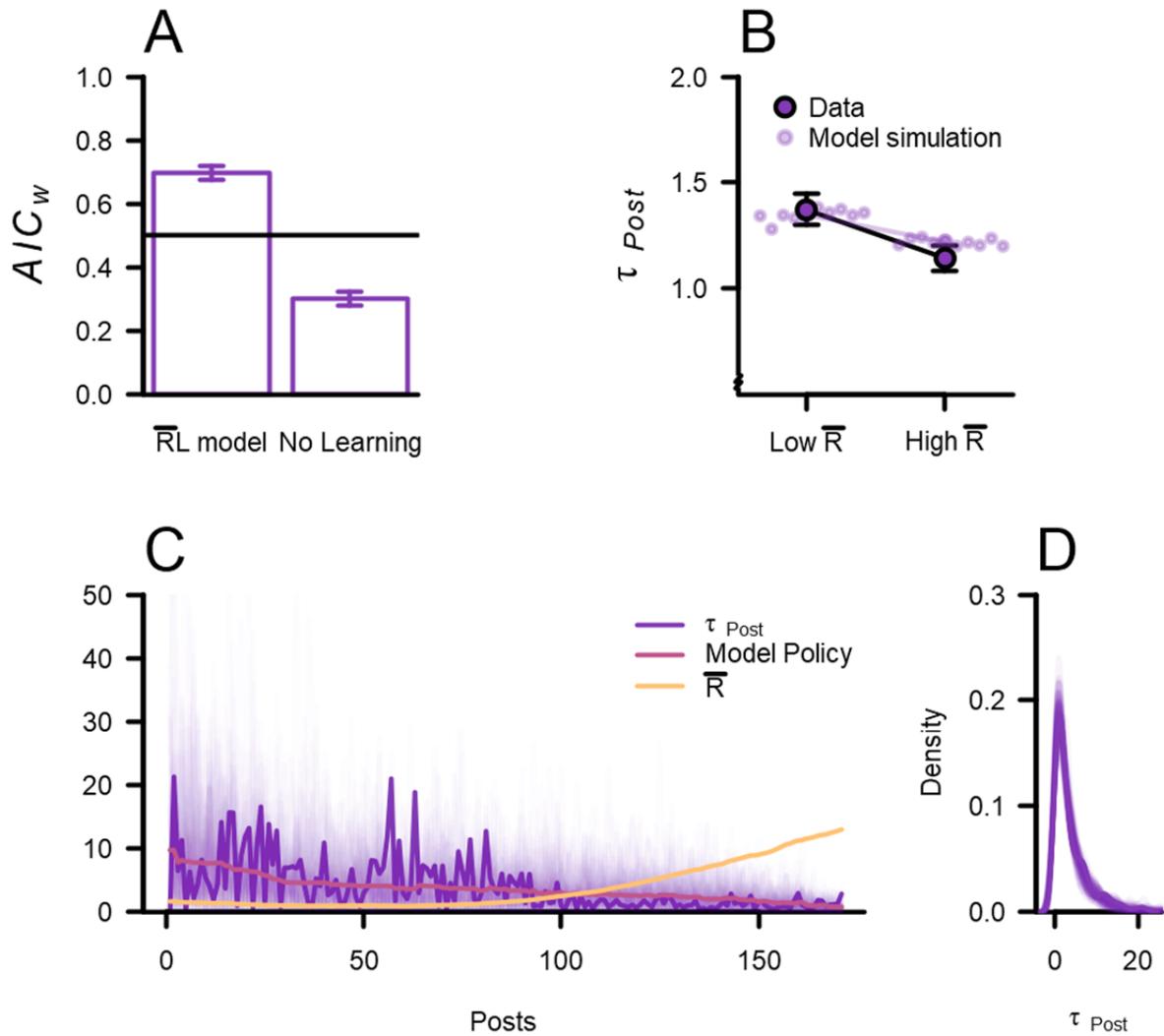


Figure 2. Behavior on Instagram is explained by reward learning (Study 1) (A) Model comparison shows that the $\bar{R}L$ model explained behavior on Instagram ($N = 2,039$) better than a model without learning. The AIC_w expresses the relative likelihood for each model. Error bars indicate the 99% CI, the horizontal line at 0.5 represents the chance level. (B) The model derived estimate of \bar{R} , the average reward rate, predicted the latency between posts. As implicated by reward learning theory, the latency between posts was shorter with high compared to low \bar{R} . Points indicate the corresponding estimates from synthetic data, based on ten generative simulation runs of the $\bar{R}L$ model (see text for details). The colored line denotes the average effect in the simulated data. Error bars indicate the 99% CI of the empirical mixed effects model estimate. (C-D) **Model fit for an example individual** (C) The posting history of an individual user over 673 days was well approximated by the $\bar{R}L$ model. The model policy (or posting threshold) denotes the average response latency predicted by the model at a given time point. The faded purple lines show 100 simulations of τ_{Post} from the estimated model policy, which illustrate the expected degree of variability given that policy, and how the empirical τ_{Post} falls within this range. The yellow line indicates the model estimate of the net reward rate, \bar{R} . Note that a higher estimated \bar{R} is associated with shorter response latencies (τ_{Post}). See Figure S2 for additional example individuals. (D) The distribution of τ_{Post} for the same individual. The faded purple line shows 100 simulations of τ_{Post} from the estimated model policy.

According to our theoretical framework, responses should be faster when the subjective reward rate is higher. Similar to how we derived model predictions (c.f., Figure 1D), we used the model-based estimate of \bar{R} (at $t-1$) dichotomized into “Low vs High” to predict the empirical τ_{Post} (at t), using log-linear mixed models (see Methods; the same conclusions are reached using continuous measures and regression models with cluster-corrected standard errors, see SI & Table S2). In support of the hypothesis that people learn to maximize social rewards, the latency between posts, τ_{Post} , was lower when \bar{R} was relatively high (Instagram ($N_{Obs} = 851946$, $N_{Users} = 2039$): $\beta = -0.18$, $SE = 0.003$, $t = -54.59$, $p < .0001$, see Figure 2B. Expressed in model comparison terms, the AIC_W for the regression model including \bar{R} was 1). Thus, increasing the average reward rate from low to high reduced average latency between posts by $\sim 18\%$, corresponding to ~ 8 hours. Providing additional support for the logic of the $\bar{R}L$ model, the effect of \bar{R} on posting latencies was stronger for individuals for whom the $\bar{R}L$ model provided a better fit (interaction Low vs High $\bar{R} * AIC_W$ [centered at 0.5]: $\beta = -0.04$, $SE = 0.008$, $t = -5.5$, $p < .0001$). In these analyses, we adjusted for $R(t-1)$ (the amount of likes received at post $t-1$), the specific post number, and the weekday of the preceding post, to ascertain the specificity of the effect of interest. We illustrate the relationship between τ_{Post} , \bar{R} , and the model policy in Figure 2 C-D, which displays the posting behavior of an example individual over a period of two years. Together, these findings indicate that online behavior conforms to principles of reward learning.

Generative model simulations. Evidence that one model explains data better than an alternative is a first step in model comparison, but better fit does not guarantee that the winning model can actually reproduce the effects of interest(34). To confirm the model results, we therefore generatively simulated the $\bar{R}L$ model (based on the median best fitting

parameters, but independent of the empirical data) and used mixed-effects models to summarize the simulations(35). Notably, the simulation makes very limited assumptions of how likes were generated (i.e., as random draws from a Poisson distribution, with identical parameters for all individuals, see SI for additional details). Nonetheless, we found that the simple reward learning (\bar{RL}) model reproduced the observed difference in response latency between high and low \bar{R} (Figure 2B). Together, our results expand the explanatory reach of learning theory from the behavior of rodents in lab experiments to the behavior of humans on social media.

Study 2

To replicate and extend the results of Study 1, we collected data from three distinct social media sites (see Methods) which, in contrast to Instagram, focus on special interest topics (Men’s fashion, Women’s fashion, Gardening, respectively). Much activity on these social media sites is focused on textual exchange rather than images, but all three contain prolific “threads”—collections of posts focused on a specific topic—with predominantly image-based content (e.g., “What are you wearing today?”, “Post pictures of your garden”), with many thousands of posts each. We limited our analyses to posts with user-generated images from such threads (see Methods and SI).

We again tested the hypothesis that social media behavior reflects social reward learning by estimating the same \bar{RL} model as in Study 1. In all three datasets (190,721 data points from 2,127 individuals), regardless of the specific topic, model comparison favored the \bar{RL} model over the model without learning (pooled mean $AIC_w = 0.77$, 99% CI [0.76,0.79], t-test against equal model likelihoods: $t(2126) = 38.84$, $p < .0001$, $x_p = 1$, see Figure 3 A-C). As in Study 1, we performed several robustness checks to verify this conclusion (see SI).

These findings converge with and generalize those of Study 1, providing platform-independent evidence that reward learning theory can help explain social media behavior.

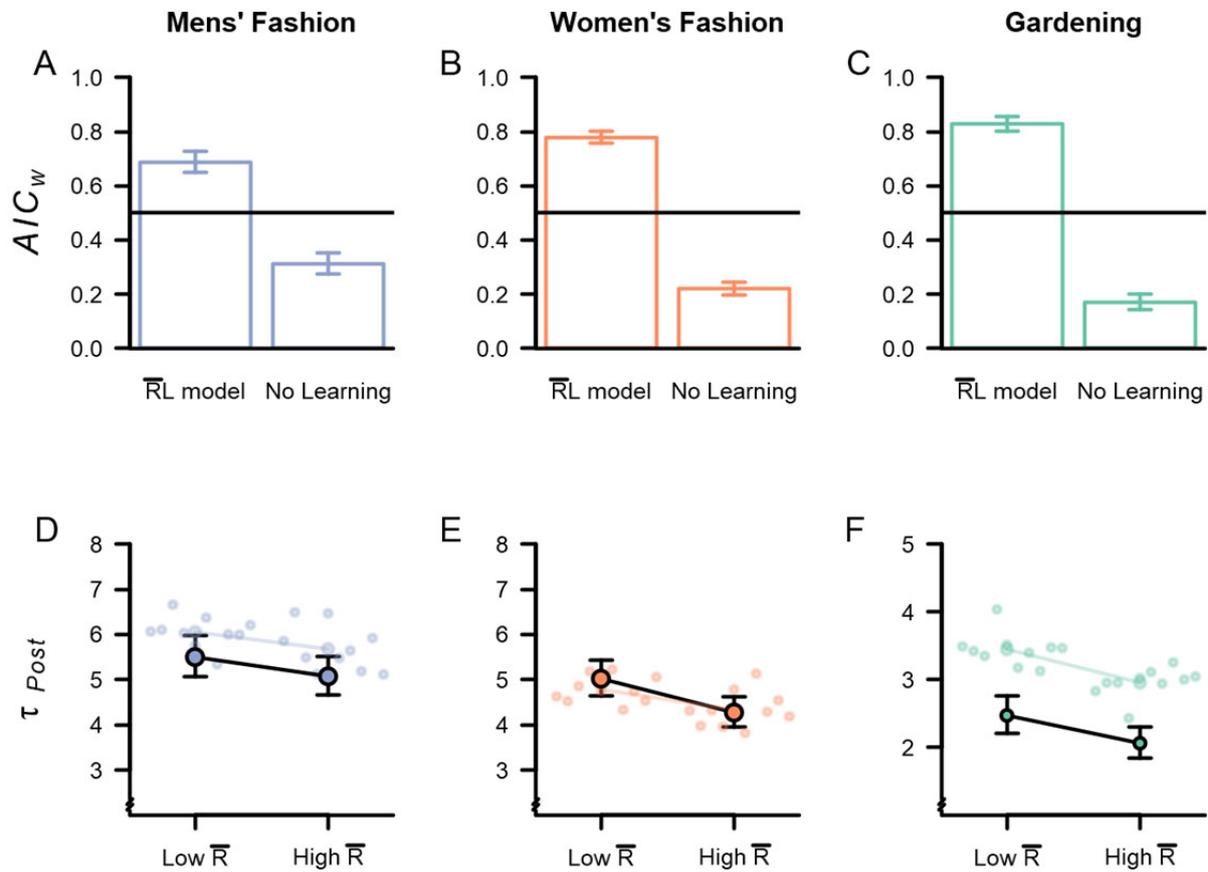


Figure 3. Signatures of reward learning on three social media sites (Study 2). (A-C) Model comparison shows that the $\bar{R}L$ model explained behavior on the three social media sites ($N = 2,127$) better than a model without learning. The AIC_w expresses the relative likelihood for each model. The exceedance probability for the $\bar{R}L$ model was 1 in all three datasets. Error bars indicate the 99% CI. **D-F)** The model derived estimate of \bar{R} , the average reward rate, predicted the latency between posts on each social media platform. In line with reward learning theory, the latency between posts was shorter with high compared to low \bar{R} . The colored points indicate the corresponding estimates from simulated data, based on ten generative simulation runs of the $\bar{R}L$ model (see text for details). The colored lines show the average effect in the simulated data. The error bars indicate the 99% CI of the empirical mixed effects model estimate.

As in Study 1, we used the model-based estimate of \bar{R} to predict the empirical τ_{Post} , using mixed models (adjusting for the same regressors as in Study 1). As expected, a higher \bar{R} predicted faster responding in all three datasets (see Figure 3D-F. Men's fashion ($N_{Obs} = 36,139$, $N_{Users} = 541$): $\beta = -0.08$, $SE = 0.016$, $t = -5.1$, $p < .0001$. Women's fashion ($N_{Obs} = 36,434$, $N_{Users} = 773$): $\beta = -0.16$, $SE = 0.02$, $t = 7.1$, $p < .0001$. Gardening: ($N_{Obs} = 118,148$, $N_{Users} = 813$): $\beta = -0.18$, $SE = 0.02$, $t = 12.09$, $p < .0001$). Thus, in these respective datasets, latencies between posts were 8%, 16%, and 18% shorter when the average reward rate was high rather than low. As in Study 1, the estimated effect of \bar{R} on posting latencies was stronger for individuals for whom the \bar{RL} model provided a better fit (see SI). Thus, regardless of platform topic, social media behavior conforms to model-based principles of reward maximization through RL.

We again conducted generative model simulations of the \bar{RL} model, based on the median estimated parameters, to test whether the model reproduced the effect of average reward rates on posting latencies. Corroborating Study 1, these simulations showed that the \bar{RL} model accurately reproduced the effect of \bar{R} observed in the data (Figure 3 D-F) (but note that the absolute level of τ_{Post} is slightly overestimated in the Gardening dataset). Together with Study 1, these results confirm that basic reward learning theory provides a powerful tool for predicting and explaining the dynamics of social media use, independent of topic.

Computational phenotyping of reward learning on social media

Having established that reward learning can explain social media behavior, we next asked whether individuals differ in the particular ways they learn from rewards on social media. To address this issue, we used the parameter estimates of the \bar{RL} model as a compact but rich description (i.e., a computational phenotype) of the mechanisms underlying behavior(36). Individual differences in these parameters can thus be viewed as differences in computational

mechanisms(36) that are interpretable across domains. For example, individual differences in learning rates have previously been linked to both genetic(37) and developmental differences(38) between individuals, while individual differences in effort cost sensitivity have been related to the dopaminergic system(39).

More specifically, we used the three parameters of the $\bar{R}L$ model, estimated for each individual from datasets 1-4 (total $n = 4,168$), as input for k-means clustering, an unsupervised method for finding sub-groups in multidimensional data. Quantitative assessment, using multiple standard criteria, showed that four clusters gave the best sub-group solution (see Figure 4A and Methods). These clusters comprised between 41% (1739 individuals) to 7% (299 individuals) of the total dataset. Importantly, although the four datasets varied in mean τ_{Post} (as reflected in the P parameter), the cluster assignment was not strongly explained by dataset (Cramér's $V = 0.3$; Cramér's V is a measure of the association between two nominal variables, where 1 denotes perfect association). This indicates that clusters captured individual differences in computational learning mechanisms, rather than idiosyncrasies of social media sites.

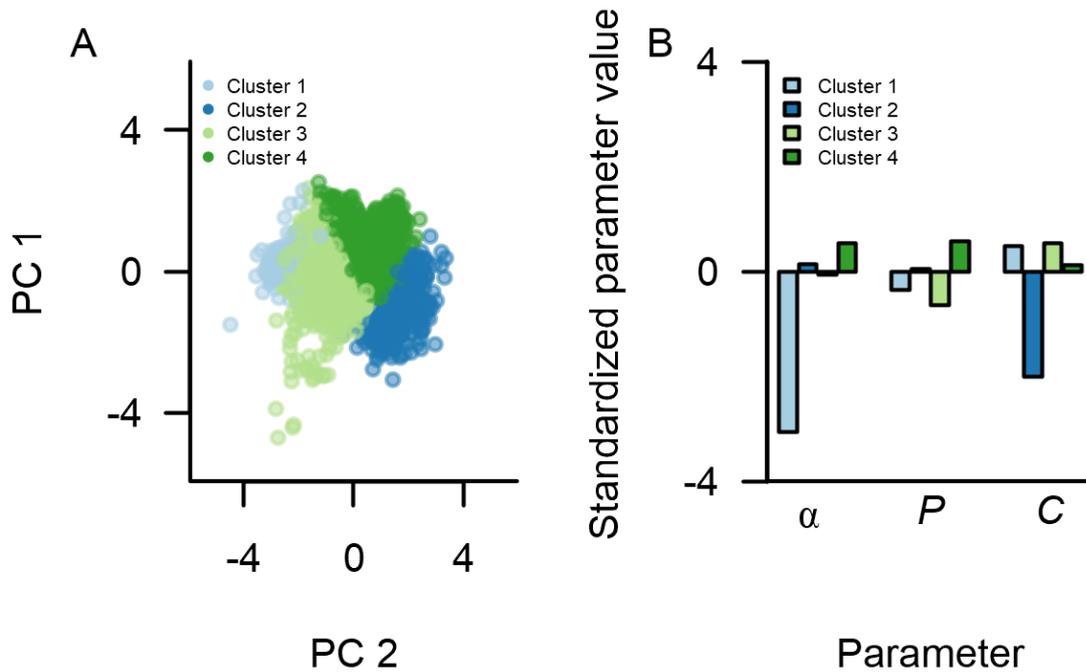


Figure 4. Computational phenotypes in reward learning on social media. (A) Cluster analysis ($n = 4,168$) of the estimated $\bar{R}L$ model parameters indicated that there were four distinct individual difference profiles in reward learning on social media. For illustration, the cluster assignments are plotted on the two first principal components (PC). (B) Profile of median $\bar{R}L$ model parameter values (standardized) for each cluster; a = learning rate, P = initial response policy, C = effort cost sensitivity.

Figure 4B illustrates the four putative computational phenotypes. For example, individuals in cluster 1 are characterized by a relatively low learning rate (a). Such individuals are especially insensitive to social rewards in their behavior (and naturally, the $\bar{R}L$ learning model provided the worst fit to these individuals relative to the model without learning: mean $AIC_w = 0.11$, vs $AIC_w \sim 0.77$ in the other three clusters). By comparison, individuals in cluster 2 are characterized by low effort cost and average learning rate, whereas cluster 4 exhibits the opposite relationship between learning rate and effort cost (and cluster 3 an intermediate profile). Individuals in both clusters 2 and 4 therefore readily post in response to social rewards, although the underlying mechanisms differ. In summary, the computational phenotyping indicates that there are important individual differences in the mechanisms underlying social media behavior.

Social comparison explains individual differences in reward learning on social media

The preceding analyses showed that people dynamically adjust their social media behavior in response to their own social rewards, as predicted by reward learning theory—a theory originally developed to test nonsocial rewards (e.g., food reward) in solitary contexts. However, given the intrinsically social context of social media use, we also expected that reward learning online would be modulated by the rewards others receive. Thus, we next asked whether individual differences in social comparison(3) might account for additional variation between individuals in how reward learning mechanisms guide social media behavior.

To answer this question, we focused on the three social media sites analyzed in Study 2, as the format of these sites facilitates direct social comparison: one’s post, and the likes it incurred, are displayed in sequential order together with others’ posts on the same forum and topic. As a model-based test for social comparison in reward learning, we modified the $\bar{R}L$ model to include an additional term, $-\xi \bar{R}_{Social}(t)$. Here, $\bar{R}_{Social}(t)$ refers to the mean number of likes per post on the forum (within the threads we analyzed, see SI) in the week preceding t , and ξ to a free parameter that determines the strength of social comparison. The social comparison term functions as a time-specific social reference level, which in practice can transform even objectively large rewards into negative prediction errors if others receive even larger rewards. This definition of social comparison is by necessity coarse, and thus conservative, as these datasets lack information regarding the specific social information to which individual users attended. However, the social comparison extension ($\xi + \bar{R}L$) of the basic model allowed us to quantitatively test if social comparison at all contributes to reward learning dynamics on social media.

Our data suggested that social comparison matters: although the $\bar{R}L$ model provided the overall best account of the data on the group level ($x_p = 1$), the $\xi + \bar{R}L$ model explained the dynamics of τ_{post} better than the $\bar{R}L$ model for a large subgroup of users (mean $AIC_W = 0.40$, across the three datasets in Study 2, see Figure 5A. Bayesian model attribution to $+\bar{R}L$: 41%). Notably, social comparison in the $\xi + \bar{R}L$ model only occurs upwards (i.e., reflecting disadvantageous inequality or envy(40)): the rewards one receives become less valuable if others receive more(17). Models that also included downward social comparison (advantageous inequality or pride/gloating(41)) provided an inferior account of the data (see SI and Table S3). The strength of the social comparison, captured by ξ , was highly variable across individuals (Figure 5B) and differed among the three social network sites: median ξ was higher on the Women’s Fashion site than on both the Men’s Fashion (Brown-Mood median test, $z = 3.23$ $p = .001$) and Gardening sites ($z = 2.91$, $p = .004$), while there was no difference among the latter two ($z = 0.13$, $p = .9$). Importantly, taking individual differences in ξ into account improved the prediction of computational phenotyping cluster assignment (because the social comparison term decomposes the effort cost term (C) of the $\bar{R}L$ model for individuals characterized by high effort cost sensitivity; see SI).

These results demonstrate that social comparison contributes to the construction of social rewards, which helps explain the dynamics of reward learning on social media. For those sensitive to social comparison, the rewards of others serve as an additional reference point for the computation of prediction errors in reward learning on social media.

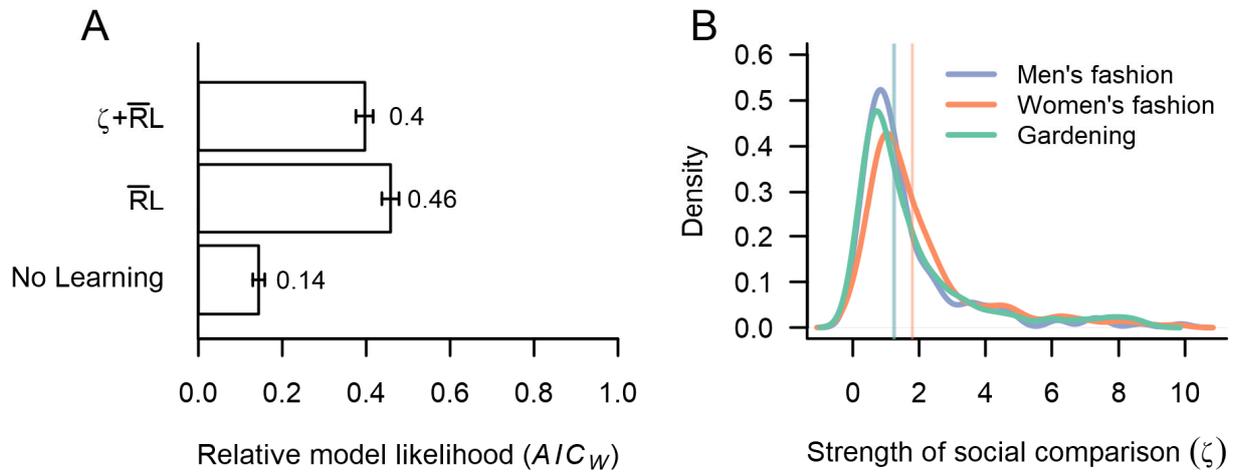


Figure 5. (A) Model comparison showed that a large minority of individuals in Study 2 was best described by the $\zeta + \bar{RL}$ model, which includes social comparison. The numbers show the exact AIC_w values. Error bars denote 99% CI. **(B) Distribution of the estimated ζ parameter.** The figure includes the subset of individuals for whom the $\zeta + \bar{RL}$ model had the highest AIC_w on the three social media sites. The vertical lines indicate the median estimated ζ parameter for each social network site.

Discussion

We investigated whether reward learning theory can help explain real-life human behavior on online social media in four large datasets. We found that social media behavior exhibited key signatures of reward learning, and that computational models inspired by RL theory, originally developed to explain the behavior of non-human animals in Skinner boxes, could quantitatively account for online behavior.

Taken together, these results advance understanding of human behavior on social media, an increasingly pervasive and profoundly consequential arena for human interaction in the 21st century. For example, it has been argued that online expressions of moral outrage, and in turn polarization(42), are fueled by social feedback, such as likes, in accordance with the principles of reward learning(43). Although we focused on the timing, rather than the content, of social media posts, our results provide clear evidence that behavior on social

media indeed follows principles of reward maximization, and that people compare their own rewards to those of others. These observations, along with its formal modeling, have broad implications for understanding and predicting multiple aspects of online behavior, including dating (e.g., learning from outcomes on dating apps), social norms, and prejudice(44).

Our computational model of social media behavior draws from RL theory originally intended to explain how non-human animals select the vigor of their responses by encoding the average net rate of rewards(26). Apart from providing a normative explanation for key behavioral regularities (e.g., the “matching law”(25)), an important aspect of this theory is the idea that \bar{R} , the average rate of reward, is encoded by the tonic, average level of dopamine(26). This idea has received some support in humans, where pharmacologically increasing the tonic level of dopamine, which according to the theory corresponds to a higher subjective reward rate, decreased average response latencies(45). Although our behavioral findings cannot speak to the neurobiological basis of reward learning on social media, the link we establish between online response latencies and the average reward rate warrant further exploration of the underlying brain mechanisms.

More generally, our results indicate that dopamine-inspired RL theory can help explain real-life individual behavior on timescales that are orders of magnitude larger than typically investigated in the lab. In turn, this insight might contribute to a more mechanistic perspective on both healthy and maladaptive (e.g. addictive(4, 5)) aspects of social media use, with the potential to inspire novel, theoretically-based design solutions or interventions. Such interventions could be individualized by applying computational phenotyping to an individual’s existing social media record (e.g., by increasing the effort cost of posting for individuals characterized by low C), thus providing novel ideographic approaches developed from theoretical models tested on large-scale data.

In conclusion, our findings reveal that basic reward learning mechanisms contribute to human behavior on social media. Understanding modern online behavior as an expression of social reward learning mechanisms offers a new window into the psychological and computational mechanisms that drive people to use social media while illuminating the link between basic, cross-species mechanisms and uniquely human modes of social interaction.

Methods

Datasets. Study 1 was based on data from a previously published study (see (27) for further information), in which data collection was based on a random sample of individuals who partook in a specific photography contest on Instagram in 2014. The dataset was fully anonymized. To allow for analyses of learning, we excluded all individuals with less than 10 posts (46). For Study 1, the final dataset consisted of 851,946 posts from 2,039 individuals. For Study 2, we obtained three datasets from three different topic-focused (Men’s fashion: *styleforum.net*, Women’s fashion: *forum.purseblog.com*, Gardening: *garden.org*, see SI for details) social media discussion forums using web scraping techniques(47) on publicly accessible data. The datasets were fully anonymized, and only included the time stamps and likes associated with posts in prolific threads focused on user-generated images (e.g., pictures of one’s clothing, see SI). For our analyses, we focused on posts with user-generated images, and, in analogy to Study 1, removed all users with fewer than ten image posts. The Study 2 dataset consisted of 190,721 posts from 2,127 individuals. This research was conducted in compliance with the US Office for Human Research Protections regulations (45 CFR 46.101 (b)).

Computational modeling. See the SI for model description and estimation methods.

Statistical analysis. All model estimation, simulations, and statistical analyses were conducted using *R*. Granger causality analysis was applied to first differenced data using the *plm* package for panel-data analysis (see SI for details)(48). Mixed effects modeling was conducted with the *lme4* package (49). All log-linear mixed effects models included a random intercept for each user. In the statistical analyses, the dependent variable τ_{Post} was log transformed (as the time between events follow an exponential distribution) to improve linearity. All predictors were standardized within individual. Degrees of freedom, test statistics, and *p*-values were derived from Satterthwaite approximations in the *lmerTest* package(50). The key statistical analyses were in addition repeated using log-linear regression models with cluster-corrected standard errors to ensure robustness (see Table S2). Prior to k-means clustering, the $\bar{R}L$ model parameter estimates were log-transformed (to improve linearity) and standardized. The optimal number of clusters was determined using the *NbClst* package (51).

Acknowledgments: We thank Andreas Olsson for helpful comments on an earlier version of the manuscript.

References

1. D. Chaffey, Global social media research summary 2019 (2019), (available at <https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>).
2. R. A. Hayes, C. T. Carr, D. Y. Wohn, One Click, Many Meanings: Interpreting Paralinguistic Digital Affordances in Social Media. *J. Broadcast. Electron. Media.* **60**, 171–187 (2016).
3. A. M. Rosenthal-von der Pütten, M. R. Hastall, S. Köcher, C. Meske, T. Heinrich, F. Labrenz, S. Ocklenburg, “Likes” as social rewards: Their role in online social comparison and decisions to like other People’s selfies. *Comput. Human Behav.* **92**, 76–86 (2019).
4. D. Kuss, M. Griffiths, Social Networking Sites and Addiction: Ten Lessons Learned. *Int. J. Environ. Res. Public Health.* **14**, 311 (2017).
5. C. S. Andreassen, Online Social Network Site Addiction: A Comprehensive Review. *Curr. Addict. Reports.* **2**, 175–184 (2015).
6. K.-Y. Lin, H.-P. Lu, Why people use social networking sites: An empirical study integrating network externalities and motivation theory. *Comput. Human Behav.* **27**, 1152–1161 (2011).

7. R. S. Sutton, A. G. Barto, *Reinforcement learning : an introduction* (MIT Press, 1998; <https://mitpress.mit.edu/books/reinforcement-learning>).
8. L. E. Sherman, A. A. Payton, L. M. Hernandez, P. M. Greenfield, M. Dapretto, The Power of the *Like* in Adolescence. *Psychol. Sci.* **27**, 1027–1035 (2016).
9. L. E. Sherman, L. M. Hernandez, P. M. Greenfield, M. Dapretto, What the brain ‘Likes’: neural correlates of providing feedback on social media. *Soc. Cogn. Affect. Neurosci.* **13**, 699–707 (2018).
10. J. Bhanji, M. Delgado, The social brain and reward: social information processing in the human striatum. *Wiley Interdiscip. Rev. ...* (2014) (available at <http://onlinelibrary.wiley.com/doi/10.1002/wcs.1266/full>).
11. R. Gu, W. Huang, J. Camilleri, P. Xu, P. Wei, S. B. Eickhoff, C. Feng, Love is analogous to money in human brain: coordinate-based and functional connectivity meta-analyses of social and monetary reward anticipation. *Neurosci. Biobehav. Rev.* (2019), doi:10.1016/J.NEUBIOREV.2019.02.017.
12. C. C. Ruff, E. Fehr, The neurobiology of rewards and values in social decision making. *Nat. Rev. Neurosci.* **15**, 549–62 (2014).
13. E. Falk, C. Scholz, Persuasion, Influence, and Value: Perspectives from Communication and Social Neuroscience. *Annu. Rev. Psychol.* **69**, 329–356 (2018).
14. A. L. Zell, L. Moeller, Are you happy for me ... on Facebook? The potential importance of “likes” and comments. *Comput. Human Behav.* **78**, 26–33 (2018).
15. D. Y. Wohn, C. T. Carr, R. A. Hayes, How Affective Is a “Like”? The Effect of Paralinguistic Digital Affordances on Perceived Social Support. *Cyberpsychology, Behav. Soc. Netw.* **19**, 562–566 (2016).
16. N. Grinberg, P. A. Dow, L. A. Adamic, M. Naaman, in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (ACM Press, New York, New York, USA, 2016; <http://dl.acm.org/citation.cfm?doid=2858036.2858501>), pp. 564–574.
17. K. Fliessbach, B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. E. Elger, A. Falk, Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum. *Science (80-.)*. **318**, 1305–1308 (2007).
18. L. Festinger, A Theory of Social Comparison Processes. *Hum. Relations.* **7**, 117–140 (1954).
19. P. N. Tobler, C. D. Fiorillo, W. Schultz, Adaptive Coding of Reward Value by Dopamine Neurons. *Science (80-.)*. **307** (2005) (available at <http://science.sciencemag.org/content/307/5715/1642>).
20. N. Grinberg, S. Kalyanaraman, L. A. Adamic, M. Naaman, in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17* (ACM Press, New York, New York, USA, 2017; <http://dl.acm.org/citation.cfm?doid=2998181.2998320>), pp. 726–739.
21. C. T. Carr, R. A. Hayes, E. M. Sumner, Predicting a Threshold of Perceived Facebook Post Success via Likes and Reactions: A Test of Explanatory Mechanisms. *Commun. Res. Reports.* **35**, 141–151 (2018).
22. J. Cheng, C. Danescu-Niculescu-Mizil, J. Leskovec, “How Community Feedback Shapes User Behavior” (2014), (available at <http://arxiv.org/abs/1405.1429>).
23. S. Das, A. Lavoie, in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems* (2014; www.ifaamas.org).
24. L. M. Hackel, D. M. Amodio, Computational neuroscience approaches to social cognition. *Curr. Opin. Psychol.* **24**, 92–97 (2018).
25. R. J. Herrnstein, On the law of effect. *J. Exp. Anal. Behav.* **13**, 243–66 (1970).
26. Y. Niv, N. D. Daw, D. Joel, P. Dayan, Tonic dopamine: opportunity costs and the

- control of response vigor. *Psychopharmacology (Berl)*. **191**, 507–520 (2007).
27. E. Ferrara, R. Interdonato, A. Tagarelli, in *Proceedings of the 25th ACM conference on Hypertext and social media - HT '14* (ACM Press, New York, New York, USA, 2014; <http://dl.acm.org/citation.cfm?doid=2631775.2631808>), pp. 24–34.
 28. C. Gerlitz, A. Helmond, The like economy: Social buttons and the data-intensive web. *New Media Soc.* **15**, 1348–1365 (2013).
 29. I. Sen, A. Aggarwal, S. Mian, S. Singh, P. Kumaraguru, A. Datta, Worth its Weight in Likes: Towards Detecting Fake Likes on Instagram (2018), doi:10.1145/3201064.3201105.
 30. I. D. Roberts, C. A. Hutcherson, Affect and Decision Making: Insights and Predictions from Computational Models. *Trends Cogn. Sci.* **23**, 602–614 (2019).
 31. Y. Niv, D. Joel, P. Dayan, A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–81 (2006).
 32. E.-J. Wagenmakers, S. Farrell, AIC model selection using Akaike weights. *Psychon. Bull. Rev.* **11**, 192–6 (2004).
 33. K. E. Stephan, W. D. Penny, J. Daunizeau, R. J. Moran, K. J. Friston, Bayesian model selection for group studies. *Neuroimage*. **46**, 1004–1017 (2009).
 34. S. Palminteri, V. Wyart, E. Koechlin, The Importance of Falsification in Computational Cognitive Modeling. *Trends Cogn. Sci.* **21**, 425–433 (2017).
 35. B. Lindström, P. N. Tobler, Incidental ostracism emerges from simple learning mechanisms. *Nat. Hum. Behav.* (2018).
 36. E. H. Patzelt, C. A. Hartley, S. J. Gershman, Computational Phenotyping: Using Models to Understand Individual Differences in Personality, Development, and Mental Illness. *Personal. Neurosci.* **1**, e18 (2018).
 37. M. J. Frank, A. a Moustafa, H. M. Haughey, T. Curran, K. E. Hutchison, Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16311–6 (2007).
 38. W. van den Bos, M. X. Cohen, T. Kahnt, E. A. Crone, Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cereb. Cortex*. **22**, 1247–55 (2012).
 39. R. Cools, The costs and benefits of brain dopamine for cognitive control. *Wiley Interdiscip. Rev. Cogn. Sci.* **7**, 317–329 (2016).
 40. E. Fehr, K. M. Schmidt, A Theory of Fairness, Competition, and Cooperation. *Q. J. Econ.* **114**, 817–868 (1999).
 41. T. A. Wills, Downward comparison principles in social psychology. *Psychol. Bull.* **90**, 245–271 (1981).
 42. W. J. Brady, J. A. Wills, J. T. Jost, J. A. Tucker, J. J. Van Bavel, Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci.* **114**, 7313–7318 (2017).
 43. M. J. Crockett, Moral outrage in the digital age. *Nat. Hum. Behav.* (2017), doi:10.1038/s41562-017-0213-3.
 44. D. M. Amodio, Social Cognition 2.0: An Interactive Memory Systems Account. *Trends Cogn. Sci.* **23**, 21–33 (2019).
 45. U. Beierholm, M. Guitart-Masip, M. Economides, R. Chowdhury, E. Düzel, R. Dolan, P. Dayan, Dopamine modulates reward-related vigor. *Neuropsychopharmacology*. **38**, 1495–503 (2013).
 46. E. Schulz, R. Bhui, B. C. Love, B. Brier, M. T. Todd, S. J. Gershman, Structured, uncertainty-driven exploration in real-world consumer choice, doi:10.1073/pnas.1821028116.
 47. R. N. Landers, R. C. Brusso, K. J. Cavanaugh, A. B. Collmus, A primer on theory-

- driven web scraping: Automatic extraction of big data from the Internet for use in psychological research. *Psychol. Methods*. **21**, 475–492 (2016).
48. Y. Croissant, G. Millo, Panel Data Econometrics in R : The **plm** Package. *J. Stat. Softw.* **27**, 1–43 (2008).
 49. D. Bates, D. Sarkar, lme4: Linear mixed-effects models using S4 classes (2007) (available at <http://www.citeulike.org/group/220/article/1080437>).
 50. A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, **lmerTest** Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**, 1–26 (2017).
 51. M. Charrad, N. Ghazzali, V. Boiteau, A. Niknafs, **NbClust** : An R Package for Determining the Relevant Number of Clusters in a Data Set. *J. Stat. Softw.* **61**, 1–36 (2014).