Gains and Losses affect Learning Differentially at Low and High Attentional Load

Kianoush Banaie Boroujeni[1], **Marcus Watson**[2], **Thilo Womelsdorf**[1,3]

[1]Department of Psychology, Vanderbilt University, Nashville, TN 37240.

[2]Department of Biology, Centre for Vision Research, York University, 4700 Keele Street, Toronto, Ontario M6J 1P3, Canada.

[3]Department of Biomedical Engineering, Vanderbilt University, Nashville, TN 37240.

Abstract

Prospective gains and losses influence cognitive processing, but it is unresolved how they modulate flexible learning in changing environments. The prospect of gains might enhance flexible learning through prioritized processing of reward-predicting stimuli, but it is unclear how far this learning benefit extends when task demands increase. Similarly, experiencing losses might facilitate learning when they trigger attentional re-orienting away from loss-inducing stimuli, but losses may also impair learning by increasing motivational costs or when negative outcomes are overgeneralized. To clarify these divergent views, we tested how varying magnitudes of gains and losses affect the flexible learning of feature values in environments that varied attentional load by increasing the number of interfering object features. With this task design we found that larger prospective gains improved learning efficacy and learning speed, but only when attentional load was low. In contrast, expecting losses impaired learning efficacy and this impairment was larger at higher attentional load. These findings functionally dissociate the contributions of gains and losses on flexible learning, suggesting they operate via separate control mechanisms. One mechanism is triggered by experiencing loss and reduces the ability to reduce distractor interference, impairs assigning credit to specific loss-inducing features and decreases efficient exploration during learning. The second mechanism is triggered by experiencing gains which enhances prioritizing reward-predicting stimulus features as long as the interference of distracting features is limited. Taken together, these results support a rational theory of cognitive control during learning suggesting that experiencing losses and experiencing distractor interference impose costs for learning.

Introduction

Anticipating gains or losses have been shown to enhance the motivational saliency of information (Berridge & Robinson, 2016; Failing & Theeuwes, 2018; Yechiam & Hochman, 2013b). Enhanced motivational saliency can be beneficial when learning about the behavioral relevance of visual objects. Learning which objects lead to higher gains should enhance the likelihood choosing those objects in the future to maximize rewards, while learning which objects lead to loss should enhance avoiding those objects in the future to minimize loss (Collins & Frank, 2014; Maia, 2010). While these scenarios are plausible from a rational point of view, empirical evidence suggests more complex consequences of gains and losses for adaptive learning.

Prospective gains are generally believed to facilitate learning and attention to reward predictive stimuli. But most available evidence is based on tasks using simple stimuli, leaving open whether the benefit of gains generalizes to settings with more complex multidimensional objects that have high demands on cognitive control. With regard to losses, there is conflicting evidence with some studies showing benefits and others showing deterioration of performance when subjects experience or anticipate losses (see below). It is not clear whether these conflicting effects of loss are due to a u-shaped dependence of loss effects on performance with only intermediate levels having positive effects (Yechiam, Ashby, & Hochman, 2019; Yechiam & Hochman, 2013a), or whether experiencing loss might lead to a generalized re-orientation away from loss-inducing situations which in some task contexts impairs the encoding of the precise features of the loss-inducing event (Barbaro, Peelen, & Hickey, 2017; Laufer, Israeli, & Paz, 2016; McTeague, Gruss, & Keil, 2015).

Distinguishing these possible effects of gains and losses on flexible learning therefore requires experimental designs that vary the complexity of the task in addition to varying the amount of gains and losses. Such an experimental approach would allow to possibly discern an inverted u-shaped effect of the magnitude of gains and losses on learning, while also clarifying limitations of gains and losses in supporting flexible learning at higher levels of task complexity. Here, we propose such an experiment to identify how gains and losses improve or impair flexible learning at systematically increasing attentional demands.

It is widely believed that prospective gains improve attention to reward predicting stimuli, which predicts that learning about stimuli should be facilitated by increasing their prospective gains. According to this view anticipating gains acts as an independent 'top-down' mechanism for attention, which has been variably described as *value-based attentional guidance* (Anderson, 2019; Bourgeois, Chelazzi, & Vuilleumier, 2016; Theeuwes, 2019; Wolfe & Horowitz, 2017), *attention for liking* (Gottlieb, 2012; Hogarth, Dickinson, & Duka, 2010), or *attention for reward* (San Martin, Appelbaum, Huettel, &

Woldorff, 2016). These attention frameworks suggest that stimuli are processed quicker and with higher accuracy when they become associated with positive outcomes (Barbaro et al., 2017; Hickey, Kaiser, & Peelen, 2015; Schacht, Adler, Chen, Guo, & Sommer, 2012). The effect of anticipating gains can be adaptive when the gain-associated stimulus is a target for goal-directed behavior, but it can also deteriorate performance when the gain-associated stimulus is distracting or task irrelevant in which case it attracts attentional resources away from more relevant stimuli (Chelazzi, Marini, Pascucci, & Turatto, 2019; Noonan, Crittenden, Jensen, & Stokes, 2018).

Compared to gains, the behavioral consequences of experiencing or anticipating loss are often described in affective and motivational terms rather than in terms of attentional facilitation or impediment (Dunsmoor & Paz, 2015). An exception to this is the so-called '*loss attention*' framework that describes how losses trigger shifts in attention to alternative options and thereby improve learning outcomes (Yechiam & Hochman, 2013a, 2014). The *loss attention* framework is based on the finding that experiencing loss causes a vigilance effect that triggers enhanced exploration of alternative options (Lejarraga & Hertwig, 2017; Yechiam & Hochman, 2013a, 2013b, 2014). The switching to alternative options following aversive loss events can be observed even when the expected values of the available alternatives are controlled for and the switching away is not explained away by an affective loss aversion response as subject with higher or lower loss aversion both show loss-induced exploration (Lejarraga & Hertwig, 2017). According to these insights, experiencing loss should improve adaptive behavior by facilitating avoidance of bad options. Consistent with this view humans and monkeys have been shown to avoid looking at visual objects paired with unpleasant consequences (such as a bitter taste or a monetary loss) (Ghazizadeh, Griggs, & Hikosaka, 2016a; Raymond & O'Brien, 2009; Schomaker, Walper, Wittmann, & Einhauser, 2017).

What is unclear, however, is whether loss-triggered shifts of attention away from a stimulus reflects an unspecific re-orienting away from a loss evoking situation or whether it also affects the precise encoding of the loss-inducing stimulus. The empirical evidence about this question is contradictory with some studies reporting better encoding and memory for loss-evoking stimuli, while other studies reporting poorer memory and insights about the precise stimuli that triggered the aversive outcomes. Evidence of a stronger memory representation of aversive outcomes comes from studies reporting increased detection speed of stimuli linked to threat-related aversive outcomes such as electric shocks (Ahs, Miller, Gordon, & Lundstrom, 2013; Li, Howard, Parrish, & Gottfried, 2008), painful sounds (McTeague et al., 2015; Rhodes, Ruiz, Rios, Nguyen, & Miskovic, 2018), threat-evoking air puffs (Ghazizadeh, Griggs, & Hikosaka, 2016b), or fear-evoking images (Ohman, Flykt, & Esteves, 2001). In these studies, subjects were faster or more accurate in responding to stimuli that were associated with aversive outcomes. The improved responding to these threat-related, aversive stimuli indicates that those stimuli are better represented than neutral stimuli and hence can guide adaptive behavior away from them. Notably, such a benefit is not restricted to threat-related aversive stimuli but can also be seen in faster response times to stimuli associated with the loss of money, which is a secondary 'learned' reward (Bucker & Theeuwes, 2016; Small et al., 2005; Suarez-Suarez, Holguin, Cadaueira, Nobre, & Doallo, 2019). For example, in an object-in-scene learning task attentional orienting to

the incorrect location was faster when subjects lost money for the object at that location in prior encounters compared to a neutral or positive outcome (Doallo, Patai, & Nobre, 2013; Suarez-Suarez et al., 2019). Similarly, when subjects are required to discriminate a peripherally presented target object they detect the stimulus faster following a short (20ms) spatial pre-cue when the cued stimulus is linked to monetary loss (Bucker & Theeuwes, 2016). This faster detection was similar for monetary gains indicating that gains and losses can have near-symmetric, beneficial effects on attentional capture. A similar benefit for loss- as well as gain- associated stimuli has also been reported when stimuli are presented briefly and subjects have to judge whether the stimulus had been presented before (O'Brien & Raymond, 2012). The discussed evidence suggests that loss-inducing stimuli have a processing advantage for rapid attentional orienting and fast perceptual decisions even when the associated loss is a secondary reward like money. However, whether this processing advantage for loss-associated stimuli can be used to improve flexible learning and adaptive behavior is unresolved.

An alternate set of studies contradicts the assumption that experiencing loss entails processing advantages by investigating not the rapid orienting away from aversive stimuli but the fine perceptual discrimination of stimuli associated with negative outcomes (Laufer et al., 2016; Laufer & Paz, 2012; Resnik, Laufer, Schechtman, Sobel, & Paz, 2011; Schechtman, Laufer, & Paz, 2010; Shalev, Paz, & Avidan, 2018). In these studies, anticipating the loss of money did not enhance but systematically reduced the processing of loss-associated stimuli, causing impaired perceptual discriminability and reduced accuracy, even when this implied losing money during the experiment (Barbaro et al., 2017; Laufer et al., 2016; Laufer & Paz, 2012; Shalev et al., 2018). For example, losing money for finding objects in natural scenes reduces the success rate of human subjects to detect those objects compared to searching for objects that promises gains (Barbaro et al., 2017). One important observation in these studies is that the detrimental effect of losses is not simply explained away by an overweighting of losses over gains, which would be suggestive of an affective loss aversion mechanism (Laufer & Paz, 2012). Rather, this literature suggests that stimuli linked to monetary loss outcomes are weaker attentional targets compared with neutral stimuli or gain-associated stimuli. This weaker representation can be found with multiple types of aversive outcomes including when stimuli are associated with monetary loss (Laufer et al., 2016; Laufer & Paz, 2012), unpleasant odors (Resnik et al., 2011), or electric shock (Shalev et al., 2018). One possible mechanism underlying this weakening of stimulus representations following aversive experience is that aversive outcomes are less precisely linked to the stimulus causing the outcome. According to this account, the credit assignment of an aversive outcome might generalize to other stimuli that share attributes with the actual stimulus whose choice caused the negative outcome. Such a generalized assignment of loss can have positive as well as negative consequences for behavioral performance (Dunsmoor & Paz, 2015; Laufer et al., 2016). It can lead to faster detection of stimuli associated with loss, including monetary loss, in situations that are similar to the original aversive situation without requiring recognizing the precise object instance that was causing the initial loss. This may lead to enhanced learning in situations in which the precise stimulus features are not critical. However, the wider generalization of loss outcomes will be detrimental in situations that require the precise encoding of the object instance that gave rise to loss.

In summary, the surveyed empirical evidence suggests a complex picture about how losses may influence adaptive behavior and flexible learning. On the one hand, experiencing or anticipating loss may enhance learning, when it triggers attention shifts away from the loss-inducing stimulus and when it enhances fast recognition of loss-inducing stimuli to more effectively avoid them. But on the other hand, evidence suggests that associating loss with a stimulus can impair learning when the task demands require precise insights about the loss-inducing stimulus features, because these features may be less well encoded after experiencing losses than gains, which will reduce their influence on behavior or attentional guidance in future encounters of the stimulus.

To understand which of these scenarios hold true, we designed a learning task that varied two main factors. First, we varied the magnitude of gains and losses to understand whether the learning effects depend on the actual gain/loss magnitudes. This factor was only rarely manipulated in the discussed studies. To achieve this, we used a token reward system in which subjects received tokens for correct and lost tokens for incorrect choices. Secondly, we varied the demands of attention (attentional load) by requiring subjects to search for the target feature in objects that varied non-rewarded, i.e. distracting features in either only one dimension (e.g. different colors), or in two or three dimensions (e.g. different colors, body shapes and body pattern) when searching for the rewarded feature. The variation of the object feature dimensionality allows testing whether losses and gains differentially facilitate or impair learning at varying attentional processing load.

With this task design we found across four rhesus monkeys first, that expecting larger gains enhanced the efficacy of learning targets at lower attentional loads, but not at the highest attentional load. Secondly, we found that experiencing loss generally decreased flexible learning, that larger losses exacerbate this effect, and that the loss induced learning impairment is worse at high attentional load.

Our study uses nonhuman primates as subjects to establish a robust animal model for understanding the influence of gains and losses on learning in cognitive tasks with translational value for humans as well as other species (for a recent review see e.g., Yee et al., 2022). Establishing this animal model will facilitate future studies about the underlying neurobiological mechanisms. Leveraging this animal model is possible because nonhuman primates readily understand a token reward/punishment system similar to humans and can track sequential gains and losses of assets before cashing them out for primary (e.g. juice) rewards (Rich & Wallis, 2017; Seo & Lee, 2009; Shidara & Richmond, 2002; Taswell, Costa, Murray, & Averbeck, 2018).

Methods

Experimental Procedures.

All animal related experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, the Society for Neuroscience Guidelines and Policies, and approved by the Vanderbilt University Institutional Animal Care and Use Committee.

Four pair-housed male macaque monkeys (8.5–14.4 kg, 6–9 years of age) were allowed separately to enter an apartment cage with a custom build, cage-mounted touchscreen Kiosk-Cognitive-Engagement Station to engage freely with the experimental task for 90–120 min per day. The Kiosk-Cognitive-Engagement Station substituted the front panel of an apartment cage with a 30 cm recessed, 21' touchscreen and a sipper tube protruding towards the monkey at a distance of ~33 cm and a height that allows animals sitting comfortably in front of the sipper tube with the touchscreen in reaching distance. Details about the Kiosk Station and the training regime are provided in (Womelsdorf et al., 2021). In brief, all animals underwent the same training regimes involving learning to touch, hold and release touch in a controlled way at all touchscreen locations. Then animals learned visual detection and discrimination of target stimuli among increasingly complex non-rewarded distracting objects, with correct choices rewarded with fluid. Objects were 3-dimensionally rendered so called Quaddles that have a parametrically controllable feature space, varying along four dimensions, including the color, body shape, arm types and surface pattern (Watson, Voloh, Naghizadeh, & Womelsdorf, 2019). Throughout training, one combination of features was never rewarded and hence termed 'neutral', which was the spherical, uniform, grey Quaddle with straight, blunt arms (Figure 1C). Relative to the features of the neutral object we could then increase the feature space for different experimental conditions by varying from trial-to-trial features from only one, two, or three dimensions relative to the neutral object. After monkeys completed the training for immediate fluid reward upon a correct choice, a token history bar was introduced and shown on the top of the monitor screen and animals performed the feature learning task for receiving tokens until they filled the token history bar with five tokens to cash out for fluid reward (Figure 1A). All animals effortlessly transitioned from immediate fluid reward delivery for correct choices to the token-based reward schedules.

The visual display, stimulus timing, reward delivery and registering of behavioral responses was controlled by the Unified Suite for Experiments (USE), which integrates an IO-controller board with a unity3D video-engine based control for displaying visual stimuli, controlling behavioral responses, and triggering reward delivery (Watson, Voloh, Thomas, Hasan, & Womelsdorf, 2019).

Task Paradigm.

The task required monkeys to learn a target feature in blocks of 35–60 trials by choosing one among three objects, each composed of multiple features (Figure 1A–C). At the beginning of each trial, a blue square with a side length of 3.5 cm (3° radius wide) appeared on the screen. To start a new trial, monkeys were required to touch and hold the square for 500 ms. Within 500 ms after touching the blue square, three stimuli appeared on the screen at 3 out of 4 possible locations with an equal distance from the screen center (10.5 cm, 17° eccentricity). Each stimulus had a diameter of 3 cm (~2.5° radius wide) and was horizontally and vertically separated from other stimuli by 15 cm (24°). To select a stimulus, monkeys were required to touch the object for a duration longer than 100 ms. If a touch was not registered within 5 seconds after the appearance of the stimuli, the trial was aborted and a new trial was started with stimuli that differed from those of the aborted trial.

Each experimental session consisted of 36 learning blocks. Four monkeys (B, F, I, and S) performed the task and completed 33/1166, 30/1080, 30/1080, and 28/1008 sessions/blocks, respectively. We used a token-reward based multi-dimensional attention task in which monkeys were required to explore the objects on the screen and determine through trial and error a target feature while learning to ignore irrelevant features and feature dimensions. Stimuli were multi-dimensional 3-D rendered Quaddles which varied in 1 to 3 features relative to a neutral Quaddle objects (Figure 1C) (Watson, Voloh, Naghizadeh, et al., 2019). The objects were 2D-viewed and had the same 3D view orientation across trials. Four different object dimensions/features were used in the task design (Shape, Color, Arm, and Pattern). For each session features from 1, 2 or 3 feature dimensions were used as potential target and distractor features. In each learning block one feature was selected to be the correct target. Attentional load was varied by increasing the number of features that varied from trial to trial to be either 1 (e.g. objects varied only in shape), 2 (e.g. objects varied in shape and patterns), or 3 (e.g. objects varied in shape, patterns and color). The target feature was un-cued and had to be searched for through trial and error in each learning block. The target feature remained the same throughout a learning block. After each block, the target feature changed, and monkeys had to explore again to find out the newly rewarded feature. The beginning of a new block was not explicitly cued but was apparent as the objects in the new block had different feature values than the previous block. Block changes were triggered randomly after 35–60 trials.

Each correct touch was followed by a yellow halo around the stimulus as visual feedback (for 500 ms), an auditory tone, and a fixed number of animated tokens traveling from the chosen object location to the token bar on top of the screen (Figure 1A). Erroneously touching a distractor object was followed by a blue halo around the touched objects, a low-pitched auditory feedback, and in the loss conditions travelling of one or three empty (grey) tokens to the token bar where the number of lost tokens were removed from the already earned tokens as an error penalty (feedback timing was the same as for correct trials). To receive a fluid reward, i.e., to cash out the tokens, monkeys had to complete 5 tokens in the token bar. When five tokens were collected the token bar flashed red/blue three times, a high pitch tone was played as auditory feedback and fluid was delivered. After cashing out the token bar was reset to five empty token placeholders. Monkeys could not go in debt, i.e. no token could be lost when there was no token on the token bar, or carry over tokens when gained tokens were more than what they needed to complete a token bar. Every block began with an empty token bar. To make sure subjects did not carry over collected tokens in the bar when there was a block change, the block change only occurred when the token bar was empty.

Experimental design.

In each learning block one fixed token reward schedule was used, randomly drawn from seven distinct schedules (see below). We simulated the token schedules to reach a nearly evenly spaced difference in reward rate while not confounding the number of positive or negative tokens with reward rate (Figure 1D,E). For simulating the reward rate with different token gains and losses we used a tangent hyperbolic function to simulate a typical learning curve by varying the number of trials needed to reach 75% accuracy from 5 to 25 trials (the

so called 'learning trials' in a block). This is the range of learning trials the subjects showed for the different attentional load conditions in 90% of blocks. We simulated 1000 blocks of learning for different combinations of gained- and lost- tokens for correct and erroneous choices, respectively. For each simulated token combination, the reward rate was calculated by dividing the frequency of the full token bar to the block length. The average reward rate was then computed as the average reward rate over all simulation runs for each token condition. The reward rate showed on average what is the probability of receiving reward per trial. Seven token conditions were designed with a combination of varying gain (G; 1, 2, 3, and 5) and loss (L; 3, 1, and 0) conditions (1G 0L, 2G −3L, 2G −1L, 2G 0L, 3G −1L, 3G 0L, and 5G 0L). The 5G/0L condition entailed providing 5 tokens for a correct choice and no tokens lost for incorrect choices. This condition was used to provide animals with more opportunity to earn fluid reward than would be possible with the conditions that have lower reward rate. Since gaining 5 tokens was immediately cashed out for fluid delivery, we do not consider this condition for the token-gain and token-loss analysis as it confounds secondary and primary reinforcement in single trials. We also did not consider for analysis those blocks in which there could be a loss of tokens (conditions with 1L or 3L), but the loss of tokens was not experienced because the subjects did not make erroneous choices that would have triggered the loss.

Overall, the subject's engagement in the task was not affected by token condition with monkeys performing all ~36 blocks that were made available to them in each daily session and requiring on average ~90 min to complete ~1200 trials. On rare occasions, a monkey took a break from the task as evident in an inter-trial interval of >1 min. in which case we excluded the affected block from the analysis. This happened in less than 1% of learning blocks. There were no significant differences between mean inter-trial intervals across different token conditions.

Analysis of learning.

The improvement of accuracy over successive trials at the beginning of a learning block reflects the learning curve, which we computed with a forward-looking 10-trial averaging window in each block (Figure 3A,C,E). We defined learning in a block as the number of trials needed to reach criterion accuracy of 75% correct trials over 10 successive trials. Monkeys on average reached learning criterion in >75% of blocks (B=67%, F=75%, I=83%, and S=74%, *see* Figure 4). We computed *post-learning accuracy* as the proportion of correct choices in all trials after the learning criterion was reached in a block (Figure 8).

Statistical analysis.

We first constructed linear mixed effect models (Pinherio & Bates, 1996) that tested how learning speed (indexed as the 'learning trial' (LT) at which criterion performance was reached) and accuracy after learning (*LT/Accuracy*) over blocks are affected by 3 factors, *attentional load* (Att_{Load}) with three levels (1, 2, and 3 distractor feature dimensions), the factor feedback gain (Fb_{Gain}) with 3 levels (Gaining tokens for correct performance: 1, 2, or 3), and the factor feedback loss (Fb_{Loss}) with 3 levels (Loss of tokens for erroneous performance: 0, −1, or −3). All three Att_{Load}, Fb_{Gain}, Fb_{Loss} factors were entered in the LMEs as continuous measure (ratio data). We additionally considered as random effects the

factor *Monkeys* with 4 levels (B, F, I, and S), and the factor *target features* (*Feat*) with 4 levels (color, pattern, arm, and shape). The random effects control for individual variations of learning and for possible biases in learning features of some dimensions better or worse than others. This LME had the form

$$LTor\ Accuracy = Att_{Load} + Fb_{Gains} + Fb_{Loss} + (1|Monkey) + (1|Feat) + b + \varepsilon \quad \text{(eq. 1).}$$

The model showed significant main effects of all three factors and random effects were inside the 95% confidence interval. To test for the interplay of motivational variables and attentional load we extended the model with the interaction effects for $Att_{Load} \times Fb_{Gain}$ and $Att_{Load} \times Fb_{Loss.}$ to:

$$LTor\ Accuracy = Att_{Load} + Fb_{Gain} + Fb_{Loss} + Att_{Load} \times Fb_{Gain} + Att_{Load} \times Fb_{Loss} + (1|Monkey) + (1|Feat) + b + \varepsilon \quad \text{(eq. 2).}$$

To compare the model with and without interaction terms we used Bayesian information criterion (BIC) as well as the Theoretical Likelihood Ratio Test (Hox, 2002) to decide which model explains the data best.

We also tested two additional models that tested whether the absolute difference of Fb_{Gain} and Fb_{Loss} played a significant role in accounting for accuracy and learning speed using ($Diff_{Gain-Loss} = Fb_{Gain} - Fb_{Loss}$) as predictor. Secondly, we tested the role of Reward Rate (*RR*) as a predictor. We calculated *RR* as the grand average of how many times a token bar was completed (reward delivery across all trials for each monkey) divided by the overall number of trials having the same attentional load and token condition. As an alternative estimation of RR we used the reward rate from our averaged simulation (Fig. 1E). LME's as formulated in eq. 2 better fitted the data than models that included the absolute difference, or either of the two estimations of *RR*, as was evident in lower Akaike Information Criterion (AIC) and BIC when these variables were not included in the models (all comparisons are summarized in Table 1). We thus do not describe these factors further.

In addition to the block-level analysis of learning and accuracy (eq. 1–2), we also investigated the trial-level to quantify how accuracy and response times (*Accuracy*/*RT*) of the monkeys over trials are modulated by four factors. The first factor was the learning status (*Learn_State*) with two levels (before and after reaching the learning criterion). In addition, we used the factor *attentional load (Att_Load)* with three levels (1, 2, and 3 distracting feature dimensions), the factor feedback gain (*Fb_Gain*) on the previous trial with three levels (gaining tokens for correct performance: 1/2/3), and the factor feedback loss on the previous trial (*Fb_Loss*) with three levels (Loss of tokens for erroneous performance: 0/−1/−3).

$$Accuracy\ or\ RT = Learn_{State} + Att_{Load} + Fb_{Gains} + Fb_{Loss} + (1|Monkey) + (1|Feat) + b + \varepsilon \quad \text{(eq. 3).}$$

Eq. 3 was also further expanded to account for interaction terms $Att_{Load} \times Fb_{Gain}$ and $Att_{Load} \times Fb_{Loss}$.

In order to quantify the effectiveness of the token bar to modulate reward expectancy we predicted accuracy or RT of the animals by how many tokens were already earned and visible in the token bar, using the factor $Token_{State}$ defined as the number tokens visible in the token-bar as formalized in eq. 4.

$$Accuracy \, or \, RT = Token_{State} + Learn_{State} + Att_{Load} + (1|Monkey) + (1|Feat) + b + \varepsilon \qquad \text{(eq. 4),}$$

We compared models with and without including the factor $Token_{State}$. We then ran 500 simulations of likelihood ratio tests and found that the alternative model which included the factor $Token_{State}$ had a better performance than the one without the factor $Token_{State}$ (p=0.009, LRstat=1073, BIC_{Token_state}= 198669, $BIC_{without \, Token\_state}$= 199730, AIC_{Token_state}= 198598, $AIC_{without \, Token\_state}$= 199670).

In separate control analyses we tested how learning varied when only conditions were considered that either only had variable gains at a fixed loss, or that had variable losses at the same, fixed gain (Figure 5). Similar to the above described models, these analyses resulted in statistically significant main effects of Fb_{Gain}, Fb_{Loss}, and attention load on learning speed and reaction time (*see* Results). Also, models with separate gain and loss feedback variables remained superior to the models with RR or $Diff_{Gain-Loss}$ (Table 1).

We also tested LMEs that excluded trials in which subjects experienced only a partial loss, i.e. when an error was committed in an experimental condition that involved the subtraction of 3 tokens (e.g. in the 2G-3L condition) but the subjects had accumulated only 2 tokens. We tested an LME model on trial-level performance accuracy but did not find differences in performance accuracy on trials following partial versus full loss.

Analyzing the interactions of motivation and attentional load.

To evaluate the influence of increasing gains and increasing losses on attentional load we calculated the Motivational Modulation Index (MMI) as a ratio scale indicating the difference in learning speed (the trials-to-criterion) in the conditions with 3 versus 1 token gains (MMI_{Gain}) or in conditions with 3 versus 0 token losses (MMI_{Loss}) relative to their sum as written in eq. 5:

$$MMI_{Gain} = \frac{LT_{G3} - LT_{G1}}{LT_{G3} + LT_{G1}} \qquad MMI_{Loss} = \frac{LT_{L3} - LT_{L0}}{LT_{L3} + LT_{L0}} \qquad \text{eq. 5}$$

Both MMI_{Loss} and MMI_{gain} were controlled for variations of attention load. For each monkey, we fitted a linear regression model (least squares), with attention load regressor, and regressed out attention load variations on learning speed (trial-to-criterion). We then calculated MMI_{Gain} and MMI_{Loss} separately for the low, medium, and high attentional load condition. We statistically tested these indices under the null hypothesis that MMI_{Gain} or

MMI_{Loss} is not different from zero by performing pairwise t-tests on each gain or loss condition given low, medium, and high attentional load. We used false discovery rate (FDR) correction of p-values for dependent samples with a significance level of 0.05 (Benjamini & Yekutieli, 2005).

To further test the load dependency of MMIs, we used a permutation approach that tested the MMI across attentional load conditions. To control for the number of blocks, we first randomly subsampled 1000 times 100 blocks from each feedback gain/loss condition and computed the MMI for each subsample separately for the feedback gain and for the feedback loss conditions in each attentional load condition (MMIs in gain conditions for 3G and 1G, and in Loss conditions for 3L and 1L). In each attentional load condition and for the feedback gain and for the feedback loss conditions we separately formed the sampling distribution of the means (1000 sample means of randomly selected 100 subsamples). We then repeated the same procedure but this time across all attentional load conditions and sampled 1000 times to form a distribution of means while controlling for equal numbers of blocks per load condition). Using bootstrapping (DiCicio & Efron, 1996), we computed the confidence intervals on the sampling distributions across all loads with an alpha level of p = 0.015 (Bonferroni corrected for family-wise error rate) under the null hypothesis that the MMI distribution for each load condition is not different from the population of all load conditions. These statistics are used in Figure's 6 and 7, for the interaction analysis of load and MMI on learning and accuracy, respectively.

Analyzing the immediate and prolonged effects of outcomes on accuracy.

To analyze the effect of token income on accuracy in both, immediate and prolonged time windows, we calculated the proportion of correct trials after a given token gained or lost on the next n^{th} upcoming trials. We computed that for high/low gains (3G and 1G) and losses (−3 L and 0L). For each nth trial, we used Wilcoxon tests to separately test whether the difference of the proportion of correct responses between high and low gains (green lines in Figure 8F) and high and low losses (red lines in Figure 8F) were significantly different from zero (separately for low and high attentional load conditions). After extracting the p-values for all 40 trials (20 trials for each load condition), we corrected the p-values by FDR-correction for dependent samples with an alpha level of 0.05. Trials significantly different from zero are denoted by red/green horizontal lines for gain/loss conditions (Figure 9F).

Results

Four monkeys performed a feature-reward learning task and collected tokens as secondary reinforcers to be cashed out for fluid reward when five tokens were collected. The task required learning a target feature in blocks of 35–60 trials through trial-and-error by choosing one among three objects composed of multiple different object features (Figure 1A,B). Attentional load was varied by increasing the number of distracting features of these objects to be either only from the same feature dimension as the rewarded target feature ('*1-Dimensional*' load), or additionally from a second feature dimension ('*2-Dimensional*' load), or from a second and third feature dimension ('*3-Dimensional*' load) (Figure 1C).

Orthogonal to attentional load, we varied between blocks the number of tokens that were gained for correct responses (1, 2, or 3 tokens), or that could be lost for erroneous responses (0, −1, −3 tokens). We selected combinations of gains and losses so that losses were used in a condition with relatively high reward rate (e.g. the condition with 3 tokens gained and 1 loss token), while other conditions had lower reward rate despite the absence of loss tokens (e.g. the condition with 1 gain and 0 loss tokens). This arrangement allowed studying the effect of losses relatively independent of overall reward rate (Figure 1D,E).

During learning, monkeys showed slower choice response times the more tokens were already earned (Figure 2A,C). This suggests that they tracked the tokens they obtained and were more careful responding the more they had earned in those trials in which they were not yet certain about the rewarded feature. After they reached the learning criterion (during plateau performance) monkeys showed faster response times the more tokens they had earned (Figure 2B,D). LME models including the variable $Token_{State}$ explained the response times better than those without *"$Token_{State}$"* variable, p=0.009, LRstat=1073, BIC_{Token_state}= 198669, $BIC_{without\ Token\_state}$= 199730, AIC_{Token_state}= 198598, $AIC_{without\ Token\_state}$= 199670) (Figure 2E).

On average, subjects completed 1080 learning blocks (SE ±32, range 1008–1166) in 30 test sessions (SE ±1, range 28–33). All monkeys showed slower learning of the target feature with increased attentional load (Figure 3 A,B) and when experiencing losing more tokens for incorrect responses (Figure 3 E,F), and all monkeys showed increased speed of learning the target feature the more tokens they could earn for correct responses (Figure 3 C,D). The same result pattern was evident for reaction times (Figure 4 A–C). The effects of load, loss, and gains were evident in significant main effects of linear mixed effects models (LME's, see Methods, $Att._{Load}$: b= 4.37, p<0.001; $feedback_{Losses}$: b=1.39, p=<0.001; $feedback_{Gains}$: b=−0.76, p=0.008). As control analyses, we compared token conditions with fixed losses (0 or −1) and variable gains (Figure 5 B,D), and fixed gains (2G or 3G) and variable losses (Figure 5 A,C). We adjusted the LMEs for both reaction time and learning speed. Similar to our previous observations we found significant main effects of gain and loss feedback on both, reaction times and learning speed (all feedback variables had main effects at p<0.001). We also tested LMEs that included as factors the absolute differences of gains and losses, and the overall reward rate (estimated as the received reward divided by the number of trials) but found that models with these factors were inferior to models without them (Table 1).

In all linear mixed effects models the factors monkeys (numbered 1–4) and the target feature dimensions (arms, body shapes, color, and surface pattern of objects) served as random grouping effects. No significant random effects were observed unless explicitly mentioned. We interpret the main effects of attentional load, token gain and token loss as reflecting changes in the efficiency of learning (Figure 3B,D,F; Fig 5A,B). Not only did monkeys learn faster at lower loads, when expecting less losses, and when expecting higher gains, but they also had fewer unlearned blocks under these same conditions (Figure 6 A–C).

The main effects of prospective gains and losses provide apparent support for a valence-specific effect of motivational incentives and penalties on attentional efficacy (Figure 7A). Because increasing losses impaired rather than enhanced learning, they are not easily

reconciled with a 'loss attention' framework that predicts that both gains and losses should similarly enhance motivational saliency and performance (Figure 7A) (Yechiam & Hochman, 2013b; Yechiam, Retzer, Telpaz, & Hochman, 2015). However, the valence-specific effect was not equally strong at low/medium/high attentional loads. While the effect of the loss magnitude interacted with attentional load (b=−3.2 p=0.012; Figure 7B), gain magnitude and attentional load did not show a significant interaction effect on learning (b=−0.13, p=0.95, LMEs; Figure 7B). We found that linear mixed effects models with interaction terms described the data better than without them (Likelihood Ratio Stat. = 60.2, p=0.009, BIC 21156 and 21184, and AIC 21091, and 21143 for [Att.$_{\text{Load}}$ × (feedback$_{\text{Losses}}$ + feedback$_{\text{gains}}$)], and [Att.$_{\text{Load}}$ + feedback$_{\text{Losses}}$ + feedback$_{\text{gains}}$], respectively). To visualize these interactions, we calculated the Motivational Modulation Index (MMI) as the difference in learning efficacy (avg. number of trials-to-criterion) when expecting to gain 3 vs. 1 token for correct choices [MMI$_{\text{Gains}}$ = Lefficacy$_{3G}$- Lefficacy$_{1G}$], and when experiencing losing 3 vs 0 token for incorrect choices [MMI$_{\text{Loss}}$ = Lefficacy$_{3L}$- Lefficacy$_{0L}$]. By calculating the MMI for each attentional load condition we can visualize whether the motivation effect of increased prospective gains and losses increased or decreased with higher attentional load (Figure 7C). We found that the detrimental effect of larger prospective losses on learning increased with attentional load, causing a larger decrease in learning efficacy at high load (Figure 7D) (permutation test, $p < 0.05$). In contrast, expecting higher gains improved learning most at low attentional load and had no measurable effect at high load (Figure 7D) (permutation test, $p < 0.05$). Pairwise t-test comparisons confirmed that MMI$_{\text{Loss}}$ was significantly different from zero (p<0.001, df = 928, tstat = −3.83; p=0.007, df = 771, tstat = −2.70; and p<0.001, df=601, tstat=−3.95; for low, medium, and high load), while MMI$_{\text{Gains}}$ was only significantly different from zero in the low load gain condition (p<0.001, df = 579, tstat = −3.39; p=0.086, df = 450, tstat = −1.72; and p=0.98, df=402, tstat=0.02, for low, medium, and high load; p-values are false discovery rate (FDR) corrected for dependent samples with an alpha level of 0.05).

The contrasting effects of gains and losses on learning efficiency were partly paralleled in post-learning accuracy (Figure 8A,B). Accuracy was enhanced with larger expected gains at lower but not at the highest attentional load conditions (t-test pairwise comparison, p<0.001, df = 579, tstat = 4.5; p=0.011, df = 450, tstat = 2.56; and p=0.76, df = 402, tstat = 0.31; for low, medium, and high load; FDR corrected for dependent samples with an alpha level of 0.05) and accuracy was decreased with larger expected losses at all loads (t-test pairwise comparison, p<0.001, df = 928, tstat=3.66; p=0.013, df = 771, tstat = 2.48; and p=0.005, df=601, tstat=2.78; for low, medium, and high load, FDR corrected for dependent samples with an alpha level of 0.05). This decrease was not modulated by load level (permutation test, p>0.05) (Figure 8A,B). In contrast to learning speed and post-learning accuracy, response times varied more symmetrically across load conditions. At low attentional load choice times were fastest with larger expected gains and with the smallest expected losses (Figure 8C,D). At medium and higher attentional loads these effects were less pronounced. All MMIs were controlled for a main effect of attentional load by regressing out attention load variations on learning speed (trial-to-criterion) (see Methods).

To understand how prospective gains and losses modulated learning efficiency on a trial-by-trial level we calculated the choice accuracy in trials immediately after experiencing a loss

of 3, 1, or 0 tokens, and after experiencing a gain of 1, 2, or 3 tokens. Experiencing the loss of 3 tokens on average led to higher performance in the subsequent trial when compared than experiencing the loss of 0 or 1 tokens (Figure 9A). This behavioral improvement after having lost 3 tokens was particularly apparent in the low attentional load condition, and less in the medium and high attentional load conditions (Figure 9B) (LME model predicting the previous trial effect on accuracy; [Att.$_{Load}$ × feedback$_{Losses}$], b=0.01, p=0.002). We quantified this effect by taking the difference in performance for losing 3 vs. 0 tokens, which confirmed that there was on average a benefit of larger penalties in the low load condition (Figure 9C). Similar to losses, experiencing larger gains improved the accuracy in the subsequent trial at low and medium attentional load, but not at high attentional load (LME model predicting previous trial effects on accuracy: for [Att.$_{Load}$ × feedback$_{Gains}$], b=−0.03, p<0.001) (Figure 9B,C). Thus, motivational improvements of post-token feedback performance adjustment were evident for token gains as well as for losses, but primarily at lower and not at higher attentional load.

Next, we analyzed how the on average improved accuracy in trials after experiencing the loss of 3 tokens (Figure 9C) might relate to reduced learning speed (Figure 7B) in the 3-Loss conditions. To test this we selected trials from the block before the learning criterion was reached and calculated accuracy in the nth trial following the experience of the token outcome using a running average window ranging from 1–20 trials. The analysis showed that after losing 3 tokens, accuracy was transiently increased compared with trials without losses (Figure 9D,E), but this effect was transient and reversed within 2 trials in the high load condition, and within 5 trials in the low load conditions (Figure 9F). For the gain conditions the results were different. Gaining 3 tokens led to a longer-lasting improvement of performance when compared to gaining 1 token. This improvement was more sustained in the low than the high attentional load condition (Figure 9E–F, the thin black lines in the upper half of the panel mark trials for which the accuracy difference of high versus low gains was significantly different from zero, Wilcoxon test, FDR corrected for dependent samples across all trials and low and high attentional load conditions with an alpha level of 0.05). These longer-lasting effects on performance closely resemble the main effects of losses and gains on the learning efficacy (Figure 7,8) and suggest that the block level learning effects be traced back to outcome-triggered performance changes at the single trial level with stronger negative effects after larger losses and stronger positive effects after larger gains.

Discussion

We found that prospective gains and losses had opposite effects on learning a relevant target feature. Experiencing losing tokens slowed learning (increased the number of trials to criterion), impaired retention (post-learning accuracy), and increased choice response times, while experiencing gaining tokens had the opposite effects. These effects varied with attentional load in opposite ways. Larger penalties for incorrect choices had maximally detrimental effects when there were many distracting features (high load). Conversely, higher gains for correct responses enhanced flexible learning at lower attentional load but had no beneficial effects at higher load. These findings were paralleled on the trial level. While there was a brief improvement of accuracy following losses for 2–5 trials

following the experience of losses during learning, accuracy declined following losses declined thereafter and on average prolonged learning. This post-error decline in learning speed was stronger with larger (3 token) loss. In contrast to losses, the experience of gains led to a more sustained improvement of accuracy in subsequent trials, consistent with better performance after gains, particularly when attentional load was low.

Together, these results document apparent asymmetries of gains and losses on learning behavior. The negative effects of losing tokens and the positive effects of gaining tokens were evident in all four monkeys. The inter-subject consistency of the token and load effects suggests that our results delineate a fundamental relationship between the motivational influences of incentives (gains) and disincentives (losses) on learning efficacy at varying attentional loads.

Experiencing loss reduces learning efficacy with increased distractor loads.

One key observation in our study is that rhesus monkeys are sensitive to visual tokens as secondary reinforcers, closely tracking their already obtained token assets (Figure 2). This finding is consistent with prior studies in rhesus monkeys showing that gaining tokens for correct performance and losing tokens for incorrect performance modulates performance in choice tasks (Rich & Wallis, 2017; Seo & Lee, 2009; Taswell et al., 2018). This sensitivity was essential in our study to functionally separate the influence of negative and positive valenced outcomes from the influence of the overall saliency of outcomes. In our study the number of tokens for correct and incorrect performance remained constant within blocks of   35 trials and thus set a reward context for learning target features among distractors (Sallet et al., 2007). When this reward context entailed losing three tokens, monkeys learned the rewarded target feature ~5–7 trials later (slower) than when incorrect choices led to no token change (depending on load, Figure 7B). At the trial level, experiencing losing 3 tokens during the initial learning briefly enhanced performance on immediately subsequent trials, suggesting subjects successfully oriented away from the loss-inducing stimulus features, but this effect reversed within three trials after the loss experience, causing a sustained decrease in performance and delayed learning when experiencing losing 3 tokens (Figure 9C,F). This result pattern suggests that experiencing losing 3 tokens led to a short-lasting re-orienting away from the loss-inducing stimuli, but it does not enhance the processing, or the remembering, of the chosen distractors. Experiencing loss rather impairs using information from the erroneously chosen objects to adjust behavior. This finding had a relatively large effect size and was rather unexpected given various previous findings that would have predicted the opposite effect. First, the *loss attention* framework suggests that experiencing a loss causes an automatic vigilance response that triggers subjects to explore alternative options other than the chosen object (Yechiam et al., 2019; Yechiam & Hochman, 2013a, 2013b). Such an enhanced exploration might have facilitated avoiding objects with features that were associated with the loss in the trials immediately following the loss experience (Figure 9C,F). But our results suggest that the short-lasting, loss-triggered re-orienting to alternative objects was not accompanied by a better encoding of the loss-inducing stimuli but predicted a weaker encoding or poorer credit assignment of the specific object features of the loss-inducing object so that the animals were less able to distinguish which objects in future trials belonged to the loss-inducing objects. Such a weaker encoding would lead to

less informed exploration after a loss, which would not facilitate but impair learning. This account is consistent with human studies showing poorer discrimination of stimuli when they are associated with monetary loss, aversive images or odors, or electric shock (Laufer et al., 2016; Laufer & Paz, 2012; Resnik et al., 2011; Schechtman et al., 2010; Shalev et al., 2018).

A second reason why the overall decrement of performance with loss was unexpected was that monkeys and humans can clearly show non-zero (>0) learning rates for negative outcomes when they are estimated separately from learning rates for positive outcomes (Caze & van der Meer, 2013; Collins & Frank, 2014; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007; Frank, Seeberger, & O'Reilly R, 2004; Seo & Lee, 2009), indicating that negative outcomes in principle are useful for improving the updating of value expectations. While we found improved performance immediately after incorrect choices, this effect disappeared after ~3 further trials and overall caused slower learning. This finding suggests that experiencing loss in a high attentional load context reduced not the learning rates per se but impaired the credit assignment process that updates the expected values of object features based on token feedback. This suggestion calls upon future investigations to clarify the nature of the loss-induced impediment using computational modeling of the specific subcomponent processes underlying the learning process (Womelsdorf, Watson, & Tiesinga, 2021).

A third reason why loss-induced impairments of learning were unexpected are prior reports that monkeys successfully learn to avoid looking at objects that are paired with negative reinforcers (such as a bitter taste) (Ghazizadeh et al., 2016a). According to this prior finding, monkeys should have effectively avoided choosing objects with loss-inducing features when encountering them again. Instead, anticipating token loss reduced their capability to avoid the objects sharing features with the object that caused token loss in previous trials, suggesting that losing a token might attract attention (and gaze) similar to threatening stimuli (like airpuffs) (Ghazizadeh et al., 2016a; White et al., 2019) and thereby causing interference that impairs avoiding the features of the loss-inducing stimuli in future trials when there were multiple features in the higher attentional load conditions.

The three discussed putative reasons for why loss might not have improved but decreased performance points to the complexity of the object space we used. When subjects lost already attained tokens for erroneously choosing an object with 1, 2, or 3 object features, they were less able to assign negative credit to a specific feature of the chosen object. Instead, they tend to over-generalize features of loss-induced objects to objects with shared features that might also contain the rewarding feature. Consequently, they did not learn from erroneous choices as efficiently as they would have learned with no or neutral feedback after errors. This account is consistent with studies showing a wider generalization of aversive outcomes and a concomitantly reduced sensitivity to the loss-inducing stimuli (Laufer et al., 2016; Laufer & Paz, 2012; Resnik et al., 2011; Schechtman et al., 2010; Shalev et al., 2018). Consistent with such reduced encoding or impaired credit assignment of the specific object features, we found variable effects on post-error adjustment (Figure 9B,D,E) and reduced longer-term performance (i.e. over ~5–20 trials) after experiencing loss (Figure 9F) with the negative effects increasing with more distractors in the higher attentional

load condition (Figure 7B,D). According to this interpretation, penalties such as losing tokens are detrimental to performance when they do not inform straightforwardly what type of features should be avoided. This view is supported by additional evidence. Firstly, when tasks have simpler, binary choice options, negative outcomes are more immediately informative about which objects should be avoided and learning from negative outcomes can be more rapid than learning from positive outcomes (Averbeck, 2017). We found a similar improvement of performance in the low load condition that lasted for 1–3 trials before performance declined below average (Figure 9F). Secondly, using aversive outcomes in a feature non-selective way might incur a survival advantage in various evolutionary meaningful settings. When a negative outcome promotes generalizing from specific aversive cues (e.g. encountering a specific predator) to a larger family of aversive events (all predator-like creatures) this can enhance fast avoidance responses in future encounters (Barbaro et al., 2017; Dunsmoor & Paz, 2015; Krypotos, Effting, Kindt, & Beckers, 2015; Laufer et al., 2016). Such a generalized response is also reminiscent of non-selective 'escape' responses that experimental animals show early during aversive learning before they transition to more cue specific avoidance responses later in learning (Maia, 2010). The outlined reasoning helps explaining why we found that experiencing loss is not helpful to avoid objects or object features when multiple, multidimensional objects define the learning environment.

Experiencing gains enhance learning efficacy but cannot compensate for distractor overload.

We found that experiencing three tokens as opposed to one token for correct choices improved the efficacy of learning relevant target features by ~4, ~1.5, and ~0 trials in the low, medium and high attentional load condition (Figure 7). On the one hand, this finding provides further quantitative support that incentives can improve learning efficacy (Berridge & Robinson, 2016; Ghazizadeh et al., 2016a; Walton & Bouret, 2019). In fact, across conditions learning was most efficient when the monkeys expected three tokens and when objects varied in only one feature dimension (1D, low attentional load condition). However, this effect disappeared at high attentional load, i.e. when objects varied trial-by-trial in features of three different dimensions (Figure 7). A reduced behavioral efficiency of incentives in light of distracting information is a known phenomenon in the problem solving field (Pink, 2009), but an unexpected finding in our task because the complexity of the actual reward rule (the rule was '*find the feature that predicts token gains*') did not vary from low to high attentional load. The only difference between these conditions was the higher load of distracting features, suggesting that the monkeys might have reached a limitation in controlling distractor interference that they could not compensate further by mobilizing additional control resources.

But what is the source of this limitation to control distractor interference? One possibility is that when attention demands increased in our task monkeys are limited in allocating sufficient control or 'mental effort' to overcome distractor interference (Hosking, Cocker, & Winstanley, 2014; Klein-Flugge, Kennerley, Friston, & Bestmann, 2016; Parvizi, Rangarajan, Shirer, Desai, & Greicius, 2013; Rudebeck, Walton, Smyth, Bannerman, & Rushworth, 2006; Shenhav et al., 2017; Walton, Bannerman, Alterescu, & Rushworth, 2003; Walton, Rudebeck, Bannerman, & Rushworth, 2007). Thus, subjects might perform

poorer at high attentional load partly because they are not allocating sufficient control of the task performance (Botvinick & Braver, 2015). Rational theories of effort and control suggest that subjects allocate control as long as the benefits of doing so outweigh the costs of exerting more control. Accordingly, subjects should perform poorer at higher demand when the benefits (e.g. the reward rate for correct performance) are not increased concomitantly (Shenhav, Botvinick, & Cohen, 2013; Shenhav et al., 2017). One strength of such a rational theory of attentional control is that it specifies the type of information that limits control, which is assumed to be the degree of cross-talk of conflicting information at high cognitive load (Shenhav et al., 2017). According to this hypothesis, effort and control demands rises concomitantly with the amount of interfering information. This view provides a parsimonious explanation for our findings at high attentional load. While incentives can increase control-allocation when there is little cross-talk of the target feature with distractor features (at low attentional load), the incentives are not able to compensate for the increased cross-talk of distracting features at the high attentional load condition. Our results therefore provide quantitative support for a rational theory of attentional control.

In our task the critical transition from sufficient to insufficient control of interference was between the medium and high attentional load condition, which corresponded to an increase of distractor features that vary trial-by-trial from 8 features (at medium attentional load) to 26 features (at high attentional load). Thus, subjects were not able to compensate for distractor interference when the number of interfering features were somewhere between 8 and 26, suggesting that prospective gains – at least when using tokens as reinforcement currency – hit a limit to enhance attentional control within this range.

Impaired learning in loss-contexts and economic decision theory.

In economic decision theory it is well documented that making choices is suboptimal in contexts that frame outcomes in terms of losses rather than gains (Tversky & Kahneman, 1981, 1991). This view from behavioral economics aims to explain which options individuals will choose in a given context, which shares some resemblance with our task, where subjects learned to choose one among three objects in learning contexts (blocks) that were framed with variable token gains and losses. In a context with higher potential losses, humans are less likely to choose an equally valued option than in a gain-context. The reason for this irrational change in choice preferences is believed to reside in an overweighting of emotional content that devalues outcomes in loss contexts (Barbaro et al., 2017; Loewenstein, Weber, Hsee, & Welch, 2001). Colloquial words for such emotional overweighting might be displeasure (Tversky & Kahneman, 1981), distress, annoyance, or frustration. Concepts behind these words may relate to primary affective responses of anger, disgust or fear. However, these more qualitative descriptors are not providing an explanatory mechanism, but rather tend to anthropomorphize the observed deterioration of learning in loss-contexts. Moreover, the economic view does not provide an explanation why the learning would be stronger affected at higher attentional load (Figure 7D).

Conclusion.

Taken together, our results document the interplay of motivational variables and attentional load during flexible learning. We first showed that learning efficacy is reduced when

attentional load is increased despite the fact that the complexity of the feature-reward target rule did not change. This finding illustrates that cognitive control processes cannot fully compensate for an increase in distractors. The failure to fully compensate for enhanced distraction was not absolute. Incentive motivation was able to enhance learning efficacy when there were distracting features of one or two feature dimensions but could not help anymore to compensate for enhanced interference when features of a third feature dimension intruded into the learning of feature values. This limitation suggests that crosstalk of distracting features represents a key process involved in cognitive effort (Shenhav et al., 2017). Moreover, the negative effect of distractor interference was exacerbated by experiencing the loss of tokens for wrong choices. This effect illustrates that negative feedback does not help to avoid loss-inducing distractor objects during learning, which documents that expecting or anticipating loss deteriorates flexible learning the relevance of objects in a multidimensional object space.

Acknowledgements

Data and code accessibility

All data supporting this study and its findings, as well as custom MATLAB code generated for analyses, are available from the corresponding author upon reasonable request.

References:

Ahs F, Miller SS, Gordon AR, & Lundstrom JN (2013). Aversive learning increases sensory detection sensitivity. Biol Psychol, 92(2), 135–141. doi:10.1016/j.biopsycho.2012.11.004 [PubMed: 23174695]

Anderson BA (2019). Neurobiology of value-driven attention. Curr Opin Psychol, 29, 27–33. doi:10.1016/j.copsyc.2018.11.004 [PubMed: 30472540]

Averbeck BB (2017). Amygdala and ventral striatum population codes implement multiple learning rates for reinforcement learning. IEEE Symposium Series on Computational Intelligence (SSCI), 1–5.

Barbaro L, Peelen MV, & Hickey C (2017). Valence, Not Utility, Underlies Reward-Driven Prioritization in Human Vision. Journal of Neuroscience, 37(43), 10438–10450. doi:10.1523/Jneurosci.1128-17.2017 [PubMed: 28951452]

Benjamini Y, & Yekutieli D (2005). False discovery rate–adjusted multiple confidence intervals for selected parameters. Journal of the American Statistical Association, 100(469), 71–81.

Berridge KC, & Robinson TE (2016). Liking, wanting, and the incentive-sensitization theory of addiction. Am Psychol, 71(8), 670–679. doi:10.1037/amp0000059 [PubMed: 27977239]

Botvinick M, & Braver T (2015). Motivation and cognitive control: from behavior to neural mechanism. Annu Rev Psychol, 66, 83–113. doi:10.1146/annurev-psych-010814-015044 [PubMed: 25251491]

Bourgeois A, Chelazzi L, & Vuilleumier P (2016). How motivation and reward learning modulate selective attention. Prog Brain Res, 229, 325–342. doi:10.1016/bs.pbr.2016.06.004 [PubMed: 27926446]

Bucker B, & Theeuwes J (2016). Appetitive and aversive outcome associations modulate exogenous cueing. Attention Perception & Psychophysics, 78(7), 2253–2265. doi:10.3758/s13414-016-1107-6

Caze RD, & van der Meer MA (2013). Adaptive properties of differential learning rates for positive and negative outcomes. Biol Cybern, 107(6), 711–719. doi:10.1007/s00422-013-0571-5 [PubMed: 24085507]

Chelazzi L, Marini F, Pascucci D, & Turatto M (2019). Getting rid of visual distractors: the why, when, how, and where. Curr Opin Psychol, 29, 135–147. doi:10.1016/j.copsyc.2019.02.004 [PubMed: 30856512]

Collins AG, & Frank MJ (2014). Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev, 121(3), 337–366. doi:10.1037/a0037015 [PubMed: 25090423]

DiCicio TJ, & Efron B (1996). Bootstrap confidence intervals. Statistical Science, 11(3), 189–228.

Doallo S, Patai EZ, & Nobre AC (2013). Reward Associations Magnify Memory-based Biases on Perception. Journal of Cognitive Neuroscience, 25(2), 245–257. doi:DOI 10.1162/jocn_a_00314 [PubMed: 23066690]

Dunsmoor JE, & Paz R (2015). Fear Generalization and Anxiety: Behavioral and Neural Mechanisms. Biological psychiatry, 78(5), 336–343. doi:10.1016/j.biopsych.2015.04.010 [PubMed: 25981173]

Failing M, & Theeuwes J (2018). Selection history: How reward modulates selectivity of visual attention. Psychon Bull Rev, 25(2), 514–538. doi:10.3758/s13423-017-1380-y [PubMed: 28986770]

Frank MJ, Moustafa AA, Haughey HM, Curran T, & Hutchison KE (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci U S A, 104(41), 16311–16316. doi:10.1073/pnas.0706111104 [PubMed: 17913879]

Frank MJ, Seeberger LC, & O'Reilly RC (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science, 306(5703), 1940–1943. doi:10.1126/science.1102941 [PubMed: 15528409]

Ghazizadeh A, Griggs W, & Hikosaka O (2016a). Ecological Origins of Object Salience: Reward, Uncertainty, Aversiveness, and Novelty. Front Neurosci, 10, 378. doi:10.3389/fnins.2016.00378 [PubMed: 27594825]

Ghazizadeh A, Griggs W, & Hikosaka O (2016b). Ecological Origins of Object Salience: Reward, Uncertainty, Aversiveness, and Novelty. Frontiers in Neuroscience, 10. doi:ARTN 378 10.3389/fnins.2016.00378

Gottlieb J (2012). Attention, learning, and the value of information. Neuron, 76(2), 281–295. doi:10.1016/j.neuron.2012.09.034 [PubMed: 23083732]

Hickey C, Kaiser D, & Peelen MV (2015). Reward Guides Attention to Object Categories in Real-World Scenes. Journal of Experimental Psychology-General, 144(2), 264–273. doi:10.1037/a0038627 [PubMed: 25559653]

Hogarth L, Dickinson A, & Duka T (2010). Selective attention to conditioned stimuli in human discrimination learning: untangling the effects of outcome prediction, valence, arousal, and uncertainty. In Mitchell CJ & L. P. ME (Eds.), Attention and associative learning (pp. 71–97). Oxford, UK: Oxford University Press.

Hosking JG, Cocker PJ, & Winstanley CA (2014). Dissociable contributions of anterior cingulate cortex and basolateral amygdala on a rodent cost/benefit decision-making task of cognitive effort. Neuropsychopharmacology, 39(7), 1558–1567. doi:10.1038/npp.2014.27 [PubMed: 24496320]

Hox J (2002). Multilevel Analysis, Techniques and Applications: Lawrence Erlbaum Associates, Inc.

Klein-Flugge MC, Kennerley SW, Friston K, & Bestmann S (2016). Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off. J Neurosci, 36(39), 10002–10015. doi:10.1523/JNEUROSCI.0292-16.2016 [PubMed: 27683898]

Krypotos AM, Effting M, Kindt M, & Beckers T (2015). Avoidance learning: a review of theoretical models and recent developments. Front Behav Neurosci, 9, 189. doi:10.3389/fnbeh.2015.00189 [PubMed: 26257618]

Laufer O, Israeli D, & Paz R (2016). Behavioral and Neural Mechanisms of Overgeneralization in Anxiety. Current Biology, 26(6), 713–722. doi:10.1016/j.cub.2016.01.023 [PubMed: 26948881]

Laufer O, & Paz R (2012). Monetary loss decreases perceptual sensitivity. Journal of Molecular Neuroscience, 48, S64–S64. Retrieved from <Go to ISI>://WOS:000310541600163

Lejarraga T, & Hertwig R (2017). How the threat of losses makes people explore more than the promise of gains. Psychonomic Bulletin & Review, 24(3), 708–720. doi:10.3758/s13423-016-1158-7 [PubMed: 27620178]

Li W, Howard JD, Parrish TB, & Gottfried JA (2008). Aversive learning enhances perceptual and cortical discrimination of indiscriminable odor cues. Science, 319(5871), 1842–1845. doi:10.1126/science.1152837 [PubMed: 18369149]

Loewenstein GF, Weber EU, Hsee CK, & Welch N (2001). Risk as feelings. Psychological Bulletin, 127(2), 267–286. doi:10.1037//0033-2909.127.2.267 [PubMed: 11316014]

Maia TV (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. Learn Behav, 38(1), 50–67. doi:10.3758/LB.38.1.50 [PubMed: 20065349]

McTeague LM, Gruss LF, & Keil A (2015). Aversive learning shapes neuronal orientation tuning in human visual cortex. Nature Communications, 6. doi:ARTN 7823 10.1038/ncomms8823

Noonan MP, Crittenden BM, Jensen O, & Stokes MG (2018). Selective inhibition of distracting input. Behav Brain Res, 355, 36–47. doi:10.1016/j.bbr.2017.10.010 [PubMed: 29042157]

O'Brien JL, & Raymond JE (2012). Learned Predictiveness Speeds Visual Processing. Psychological Science, 23(4), 359–363. doi:10.1177/0956797611429800 [PubMed: 22399415]

Ohman A, Flykt A, & Esteves F (2001). Emotion drives attention: detecting the snake in the grass. J Exp Psychol Gen, 130(3), 466–478. doi:10.1037//0096-3445.130.3.466 [PubMed: 11561921]

Parvizi J, Rangarajan V, Shirer WR, Desai N, & Greicius MD (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. Neuron, 80(6), 1359–1367. doi:10.1016/j.neuron.2013.10.057 [PubMed: 24316296]

Pinherio JC, & Bates DM (1996). Unconstrained Parametrizations for Variance-Covariance Matrices. Statistics and Computing, 6, 289–296.

Pink DH (2009). Drive: The Surprising Truth About What Motivates. New York: Riverhead Books.

Raymond JE, & O'Brien JL (2009). Selective visual attention and motivation: the consequences of value learning in an attentional blink task. Psychological Science, 20(8), 981–988. doi:10.1111/j.1467-9280.2009.02391.x [PubMed: 19549080]

Resnik J, Laufer O, Schechtman E, Sobel N, & Paz R (2011). Auditory aversive learning increases discrimination thresholds. Journal of Molecular Neuroscience, 45(Suppl 1), S96–S97. Retrieved from <Go to ISI>://WOS:000520322000251

Rhodes LJ, Ruiz A, Rios M, Nguyen T, & Miskovic V (2018). Differential aversive learning enhances orientation discrimination. Cogn Emot, 32(4), 885–891. doi:10.1080/02699931.2017.1347084 [PubMed: 28683593]

Rich EL, & Wallis JD (2017). Spatiotemporal dynamics of information encoding revealed in orbitofrontal high-gamma. Nat Commun, 8(1), 1139. doi:10.1038/s41467-017-01253-5 [PubMed: 29074960]

Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, & Rushworth MF (2006). Separate neural pathways process different decision costs. Nat Neurosci, 9(9), 1161–1168. doi:10.1038/nn1756 [PubMed: 16921368]

Sallet J, Quilodran R, Rothe M, Vezoli J, Joseph JP, & Procyk E (2007). Expectations, gains, and losses in the anterior cingulate cortex. Cogn Affect Behav Neurosci, 7(4), 327–336. doi:10.3758/cabn.7.4.327 [PubMed: 18189006]

San Martin R, Appelbaum LG, Huettel SA, & Woldorff MG (2016). Cortical Brain Activity Reflecting Attentional Biasing Toward Reward-Predicting Cues Covaries with Economic Decision-Making Performance. Cerebral Cortex, 26(1), 1–11. doi:10.1093/cercor/bhu160 [PubMed: 25139941]

Schacht A, Adler N, Chen P, Guo T, & Sommer W (2012). Association with positive outcome induces early effects in event-related brain potentials. Biol Psychol, 89(1), 130–136. doi:10.1016/j.biopsycho.2011.10.001 [PubMed: 22027086]

Schechtman E, Laufer O, & Paz R (2010). Negative Valence Widens Generalization of Learning. Journal of Neuroscience, 30(31), 10460–10464. doi:10.1523/Jneurosci.2377-10.2010 [PubMed: 20685988]

Schomaker J, Walper D, Wittmann BC, & Einhauser W (2017). Attention in natural scenes: Affective-motivational factors guide gaze independently of visual salience. Vision Research, 133, 161–175. doi:10.1016/j.visres.2017.02.003 [PubMed: 28279712]

Seo H, & Lee D (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. J Neurosci, 29(11), 3627–3641. doi:10.1523/JNEUROSCI.4726-08.2009 [PubMed: 19295166]

Shalev L, Paz R, & Avidan G (2018). Visual Aversive Learning Compromises Sensory Discrimination. Journal of Neuroscience, 38(11), 2766–2779. doi:10.1523/Jneurosci.0889-17.2017 [PubMed: 29439168]

Shenhav A, Botvinick MM, & Cohen JD (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron, 79(2), 217–240. doi:10.1016/j.neuron.2013.07.007 [PubMed: 23889930]

Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen JD, & Botvinick MM (2017). Toward a Rational and Mechanistic Account of Mental Effort. Annu Rev Neurosci, 40, 99–124. doi:10.1146/annurev-neuro-072116-031526 [PubMed: 28375769]

Shidara M, & Richmond BJ (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. Science, 296(5573), 1709–1711. doi:10.1126/science.1069504 [PubMed: 12040201]

Small DM, Gitelman D, Simmons K, Bloise SM, Parrish T, & Mesulam MM (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. Cereb Cortex, 15(12), 1855–1865. doi:10.1093/cercor/bhi063 [PubMed: 15746002]

Suarez-Suarez S, Holguin SR, Cadaueira F, Nobre AC, & Doallo S (2019). Punishment-related memory-guided attention: Neural dynamics of perceptual modulation. Cortex, 115, 231–245. doi:10.1016/j.cortex.2019.01.029 [PubMed: 30852377]

Taswell CA, Costa VD, Murray EA, & Averbeck BB (2018). Ventral striatum's role in learning from gains and losses. Proc Natl Acad Sci U S A, 115(52), E12398–E12406. doi:10.1073/pnas.1809833115 [PubMed: 30545910]

Theeuwes J (2019). Goal-driven, stimulus-driven, and history-driven selection. Curr Opin Psychol, 29, 97–101. doi:10.1016/j.copsyc.2018.12.024 [PubMed: 30711911]

Tversky A, & Kahneman D (1981). The framing of decisions and the psychology of choice. Science, 211(4481), 453–458. doi:10.1126/science.7455683 [PubMed: 7455683]

Tversky A, & Kahneman D (1991). Loss Aversion in Riskless Choice - a Reference-Dependent Model. Quarterly Journal of Economics, 106(4), 1039–1061. doi:Doi 10.2307/2937956

Walton ME, Bannerman DM, Alterescu K, & Rushworth MF (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. J Neurosci, 23(16), 6475–6479. [PubMed: 12878688]

Walton ME, & Bouret S (2019). What Is the Relationship between Dopamine and Effort? Trends Neurosci, 42(2), 79–91. doi:10.1016/j.tins.2018.10.001 [PubMed: 30391016]

Walton ME, Rudebeck PH, Bannerman DM, & Rushworth MF (2007). Calculating the cost of acting in frontal cortex. Ann N Y Acad Sci, 1104, 340–356. doi:10.1196/annals.1390.009 [PubMed: 17360802]

Watson MR, Voloh B, Naghizadeh M, & Womelsdorf T (2019). Quaddles: A multidimensional 3-D object set with parametrically controlled and customizable features. Behav Res Methods, 51(6), 2522–2532. doi:10.3758/s13428-018-1097-5 [PubMed: 30088255]

Watson MR, Voloh B, Thomas C, Hasan A, & Womelsdorf T (2019). USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents. Journal of Neuroscience Methods, 1–35. doi:10.1101/434944

White JK, Bromberg-Martin ES, Heilbronner SR, Zhang K, Pai J, Haber SN, & Monosov IE (2019). A neural network for information seeking. Nat Commun, 10(1), 5168. doi:10.1038/s41467-019-13135-z [PubMed: 31727893]

Wolfe JM, & Horowitz TS (2017). Five factors that guide attention in visual search. Nature Human Behavior, 1, 1–8.

Womelsdorf T, Thomas C, Neumann A, Watson MR, Banaie Boroujeni K, Hassani SA, Parker J, Hoffman KL (2021). A Kiosk Station for the Assessment of Multiple Cognitive Domains and Cognitive Enrichment of Monkeys. Frontiers in Behavioral Neuroscience.

Womelsdorf T, Watson MR, and Tiesinga P (2021). Learning at Variable Attentional Load Requires Cooperation of Working Memory, Meta-learning and Attention-augmented Reinforcement Learning. Journal of Cognitive Neuroscience 1–29.

Yechiam E, Ashby NJS, & Hochman G (2019). Are We Attracted by Losses? Boundary Conditions for the Approach and Avoidance Effects of Losses. Journal of Experimental Psychology-Learning Memory and Cognition, 45(4), 591–605. doi:10.1037/xlm0000607 [PubMed: 29999403]

Yechiam E, & Hochman G (2013a). Loss-aversion or loss-attention: The impact of losses on cognitive performance. Cognitive Psychology, 66(2), 212–231. doi:10.1016/j.cogpsych.2012.12.001 [PubMed: 23334108]

Yechiam E, & Hochman G (2013b). Losses as Modulators of Attention: Review and Analysis of the Unique Effects of Losses Over Gains. Psychological Bulletin, 139(2), 497–518. doi:10.1037/a0029383 [PubMed: 22823738]

Yechiam E, & Hochman G (2014). Loss Attention in a Dual-Task Setting. Psychological Science, 25(2), 494–502. doi:10.1177/0956797613510725 [PubMed: 24357614]

Yechiam E, Retzer M, Telpaz A, & Hochman G (2015). Losses as ecological guides: Minor losses lead to maximization and not to avoidance. Cognition, 139, 10–17. doi:10.1016/j.cognition.2015.03.001 [PubMed: 25797454]

Yee DM, Leng X, Shenhav A, Braver TS (2022). Aversive motivation and cognitive control. Neuroscience & Biobehavioral Reviews 133:104493. [PubMed: 34910931]
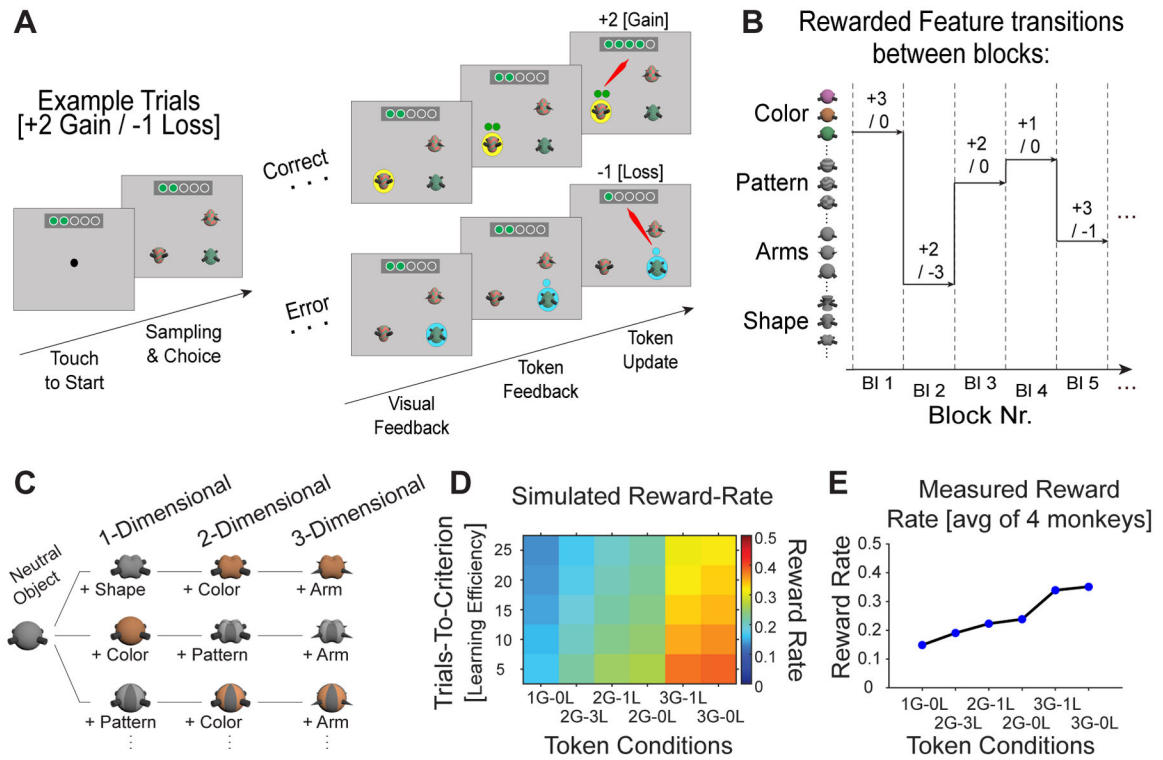
Figure 1. Task paradigm varying attentional load and token gains and losses.
(A) The trial sequence starts with presenting 3 objects. The monkey choses one by touching it. Following a correct choice a yellow halo provides visual feedback, then green tokens are presented above the stimulus for 0.3s before they are animated to move upwards towards the token bar where the gained tokens are added. Following an error trial visual feedback is cyan, then empty token(s) indicate the amount of tokens that will be lost (here: −1 token). The token update moves the empty token to the token bar where green tokens are removed from the token bar. When 5 tokens have been collected the token bar blinks red/blue, three drops of fluid are delivered, and the token bar reset. **(B)** Over blocks of 35–60 trials one object feature (here: in rows) is associated with either 1,2, or 3 token gains, while objects with other features are linked to either, 0, −1, or −3 token loss. **(C)** Attentional load is varied by increasing the number of features that define objects. The one-dimensional (1D) load condition presents objects that vary in only one feature dimension relative to a neutral object, the 2D load varies features of two feature dimensions, and the 3D load condition varies three feature dimensions. For example, a 3D object has different shapes, colors, and arm types across trials (but the same neutral pattern). **(D)** Simulation of the expected reward rate animals receive with different combinations of token-gains and -losses (*x-axis*) given different learning speed of the task (*y-axis*). **(E)** Actual reward rates (y-axis) for different token conditions (*x-axis*) based on their learning speed across 4 monkeys.
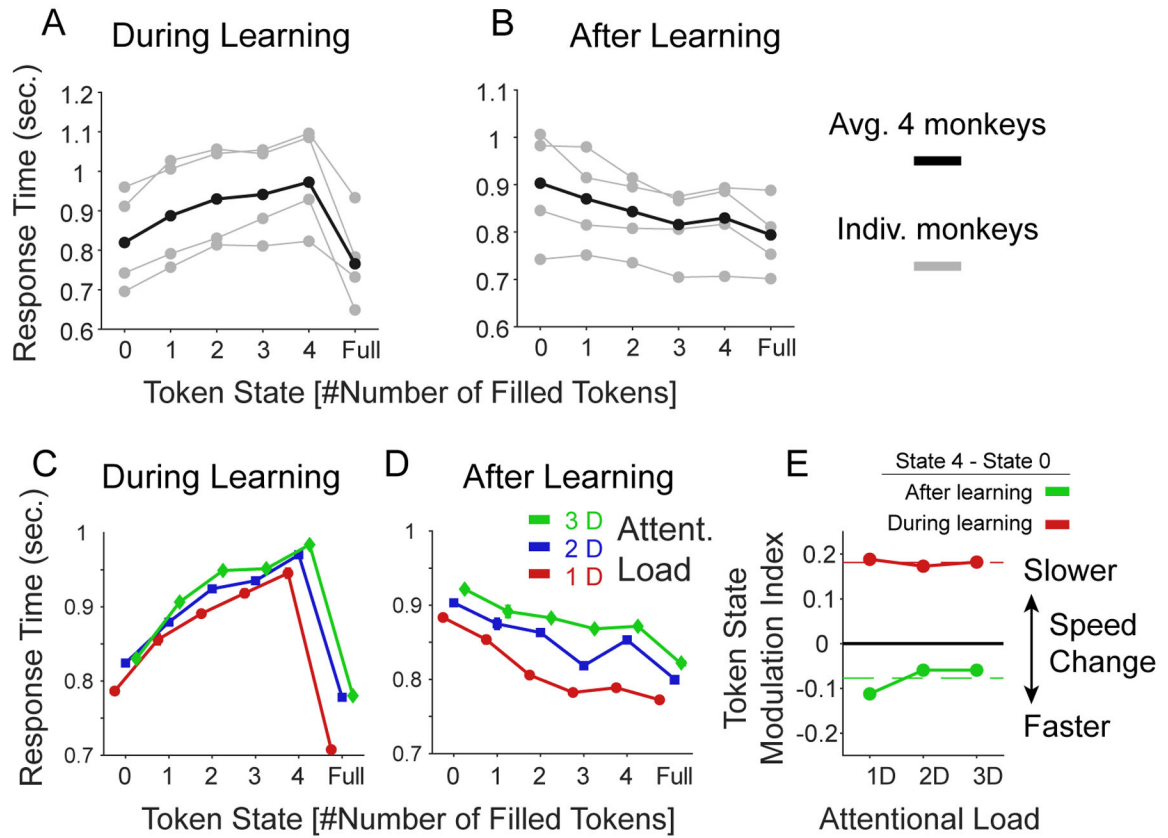
Figure 2. Effects of the number of earned tokens ('Token State') on response times.
(A) Average response times as a function of the number of earned tokens visible in the token bar (*x-axis*) for individual monkeys (grey) and their average (in black). Included were only trials prior to reaching the learning criterion. Over all trials monkeys (B, F, I, and S) showed average response time of 785, 914, 938, and 733 ms, respectively. **(B)** Same as *A* but including only trials after the learning criterion was reached. **(C-D):** Same format as *A-B* showing the average response times across monkeys for the low, medium, and high attentional load condition during learning (*C*) and after learning (E). **(E)** The Token State modulation index (*y-axis*) shows the difference in response times (RT's) when the animal had 4 tokens earned versus one token earned. During learning (*red*) RT's were slower with 4 than 1 earned tokens to similar degrees for different attentional loads (x-axis). This pattern reversed after learning was achieved (*green*). Dashed red lines show ground mean across all attention loads.
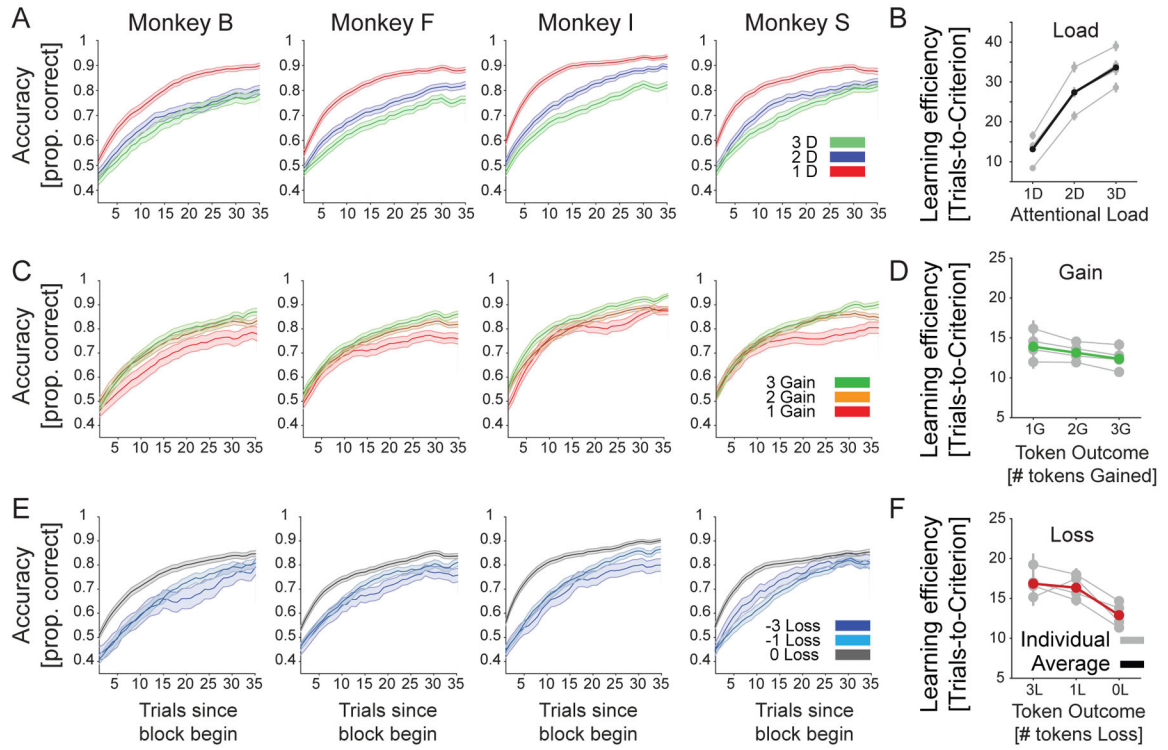
Figure 3. Average learning curves for each monkey and load, loss, and gain conditions.
(A) The proportion correct performance (*y-axis*) for low/medium/high attentional load for each of four monkeys relative to the trial since the beginning of a learning block (*x-axis*). **(B)**: The number of trials-to-reach criterion (*y-axis*) for low/medium/high attentional load (*x-axis*) for each monkey (in *grey*) and their average (in *black*). **(C)** Same as A showing the learning for blocks in which 1, 2, or 3 tokens were gained for correct performance. Red line shows average across monkeys. **(D)** Same as *B* for blocks where monkeys expected to lose 0, 1, or 3 tokens for incorrect choices. **(E)** Same as *a* and *b* showing the learning for blocks in which 0, 1, or 3 tokens were lost for incorrect performance. Green line shows average across monkeys. Errors are SE's. **(F)** Same as B for blocks where monkeys expected to win 1, 2, or 3 tokens for correct choices.
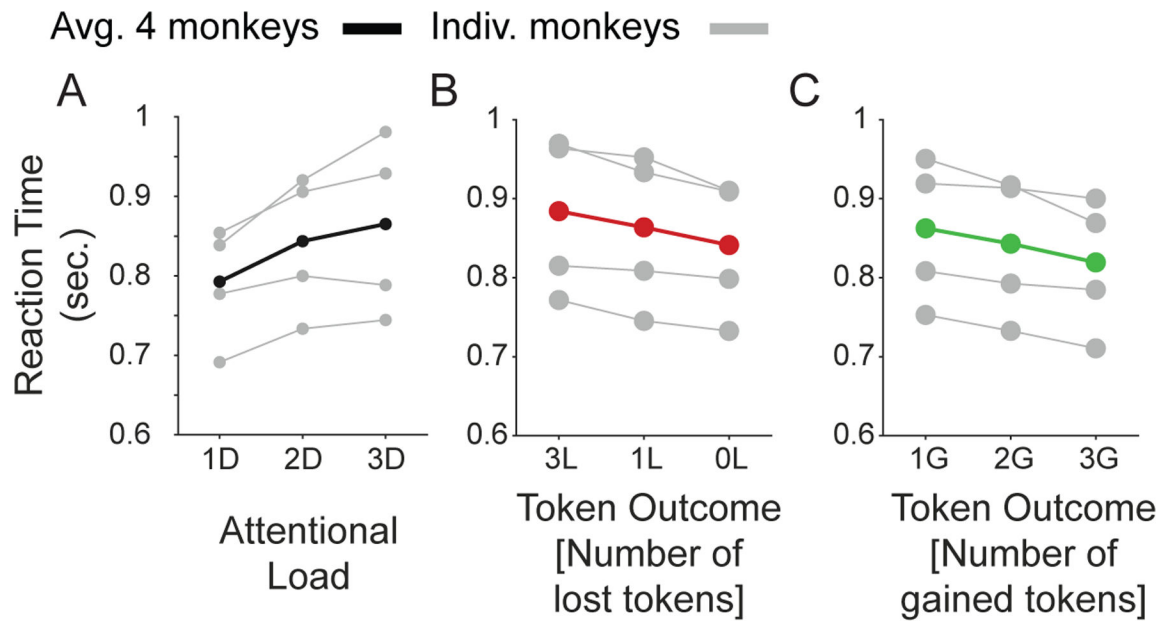
Figure 4. Main effects of attentional load, number of expected token-gains and expected token-loss on and response time.

(A) The response times (*y-axis*) for low/medium/high attentional load (*x-axis*) for each monkey (in *grey*) and their average (in *black*). **(b)** Same as *A* for blocks where monkeys expected to lose 0, 1, or 3 tokens for incorrect choices. **(c)** Same as B for blocks where monkeys expected to win 1, 2, or 3 tokens for correct choices.
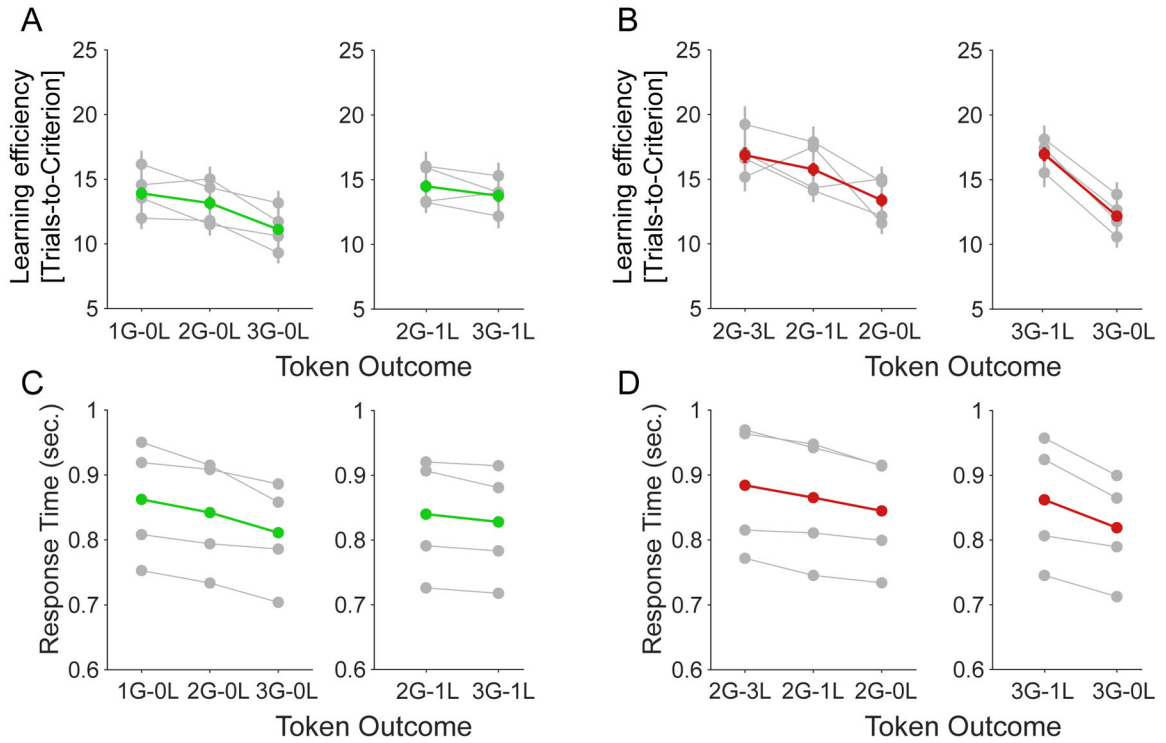
Figure 5. Controlled analyses on main effects of expected token-gains/loss when token-loss/gain held fixed on learning speed and response time.

(A) The learning speed (*y-axis*) for variable gains for fixed losses of 0 (left) and −1 (right). **(B)** The learning speed (*y-axis*) for variable losses for fixed gains of 2 (left) and 3 (right). low/medium/high attentional load (*x-axis*) for each monkey (in *grey*) and their average (in *green*). **(C)** The response time (*y-axis*) for variable gains for fixed losses of 0 (left) and −1 (right). **(D)** The response time (*y-axis*) for variable losses for fixed gains of 2 (left) and 3 (right).
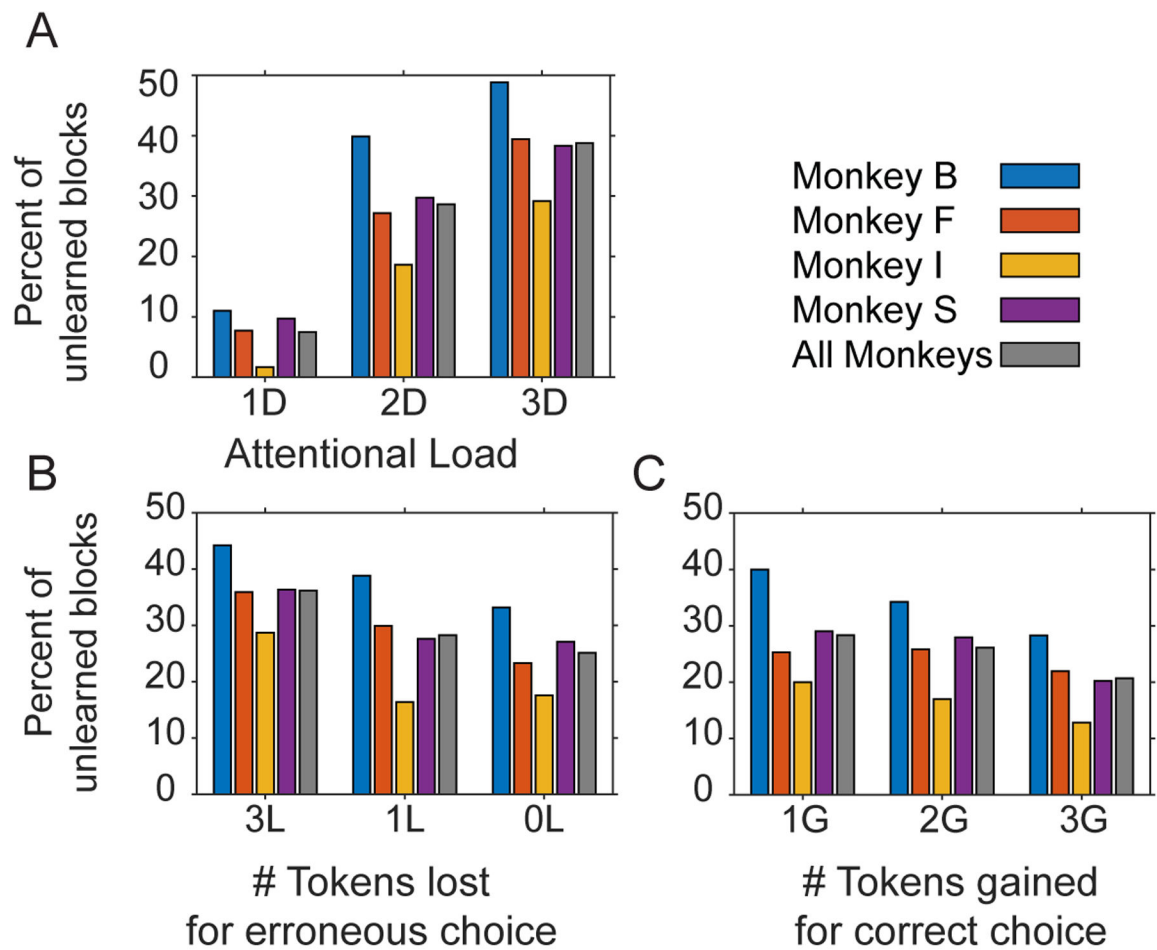
Figure 6. Proportion of unlearned blocks across conditions.
(A) The number of unlearned blocks (*y-axis*) for low/medium/high attentional load (*x-axis*) for each monkey (in *color*) and their average (in *grey*). **(B,C)** Same as (*A*) for loss (*B*) and for the gain conditions (*C*).
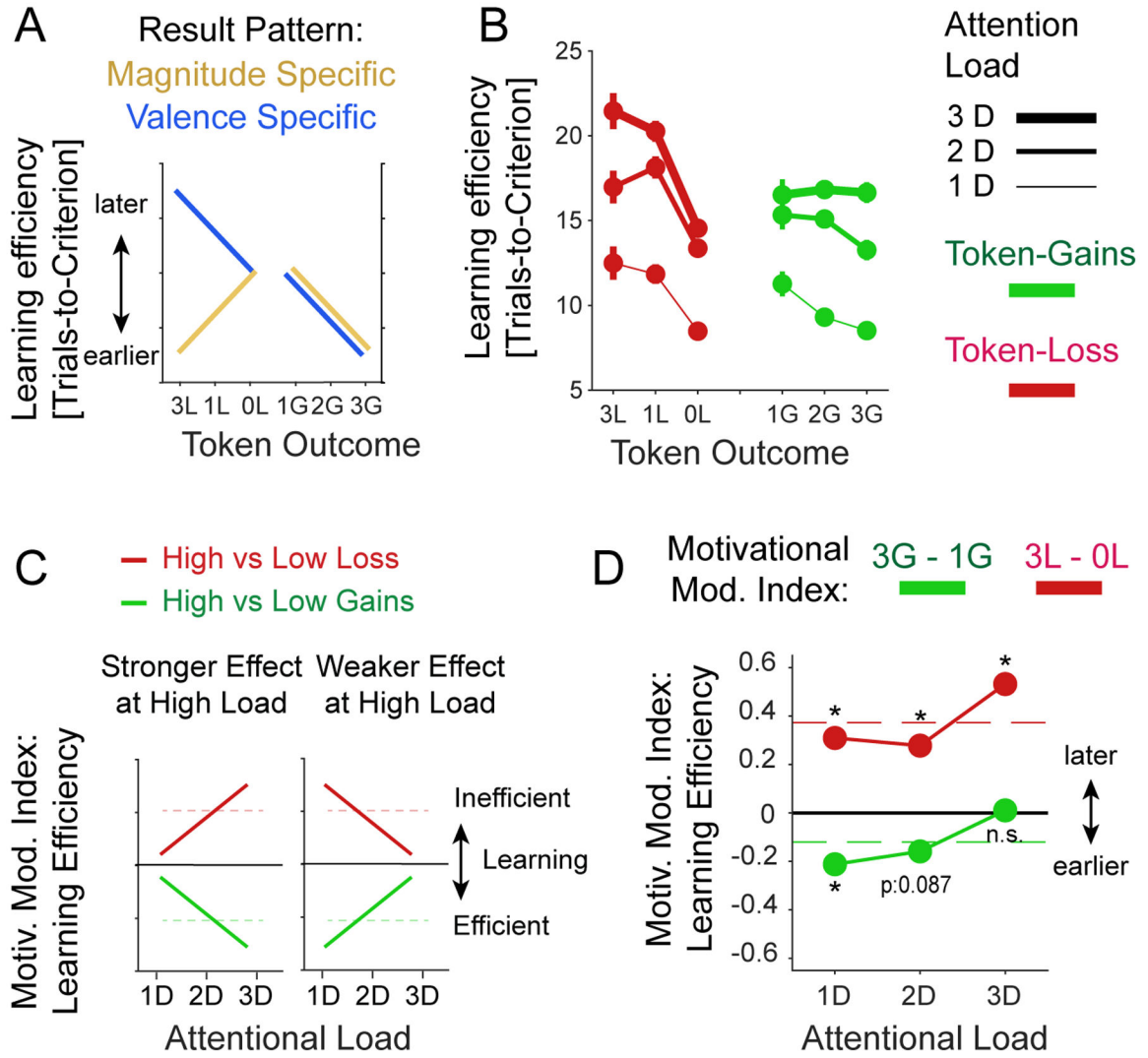
Figure 7. The effect of attentional load and expected token gain/loss on learning efficacy.
(A) A magnitude-specific hypothesis (*orange*) predicts that learning efficacy is high (learning criterion is reached early) when the absolute magnitude of expected loss (to the left) and expected gains (to the right) is high. A valence-specific hypothesis (*blue*) predicts that learning efficacy is improved with high expected gains (to the right) and decreased with larger penalties/losses. **(B)** Average learning efficacy across four monkeys at low/medium/high attentional load (line thickness) in blocks with increased expected token-loss (to the left, in red) and with increased expected token gains (to the right, green). **(C)** Hypothetical interactions of expected gains/losses and attentional load. Larger incentives/penalties might have a stronger positive/negative effect at higher load (*left*) or a weaker effect at higher load (*right*). The predictions can be quantified with the Motivational Modulation index, which is the difference of learning efficacy for the high vs. low gains conditions (or high vs low loss conditions). **(D)** Average motivational modulation index shows that the slowing effect of larger penalties increased with higher attentional load (*red*). In contrast, the enhanced learning efficacy with higher gain expectations are larger at lower attentional load and absent

at high attentional load (*green*). Dashed red lines show ground mean across all attention loads.
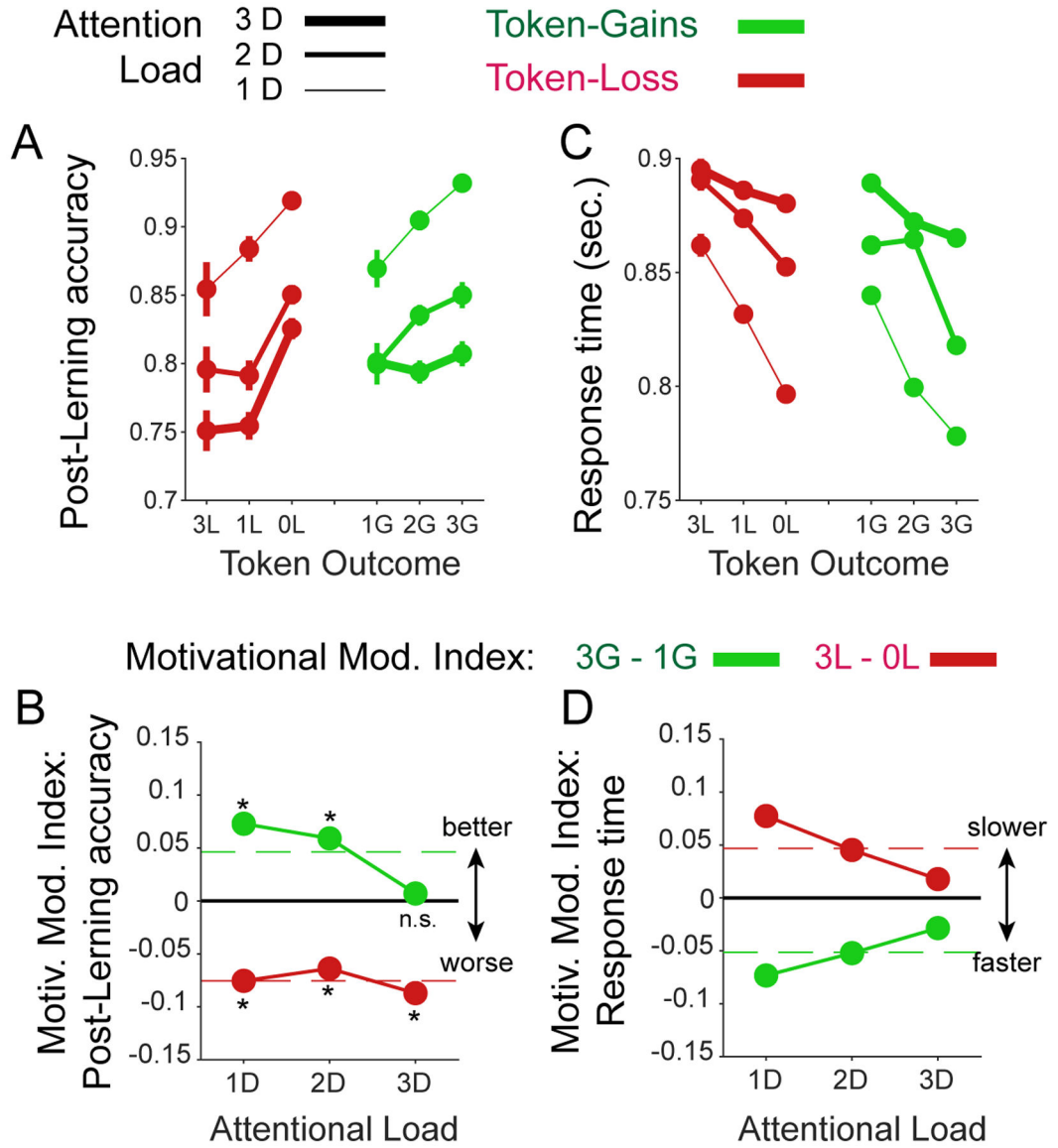
Figure 8. The effect of attentional load and token gain/loss expectancy on post-learning performance and response times.

(**A**) Post-learning accuracy (*y-axis*) when expecting varying token loss (*red*) and gains (*green*) at low/medium/high attentional load (line thickness). Overall, learning efficiency decreased with larger penalties and improved with larger expected token-gains. (**B**): The motivation modulation index for low/medium/high attentional load (*x-axis*) shows that the improvement with higher gains was absent at high load, and the detrimental effect of penalties on performance was evident at all loads. (**C,D**) Same format as A,B for response times. Subject slowed down when expecting larger penalties and speed up responses when expecting larger gains (**C**). These effects were largest at low attentional load and decreased at higher load (**D**). Dashed red lines show ground mean across all attention loads.
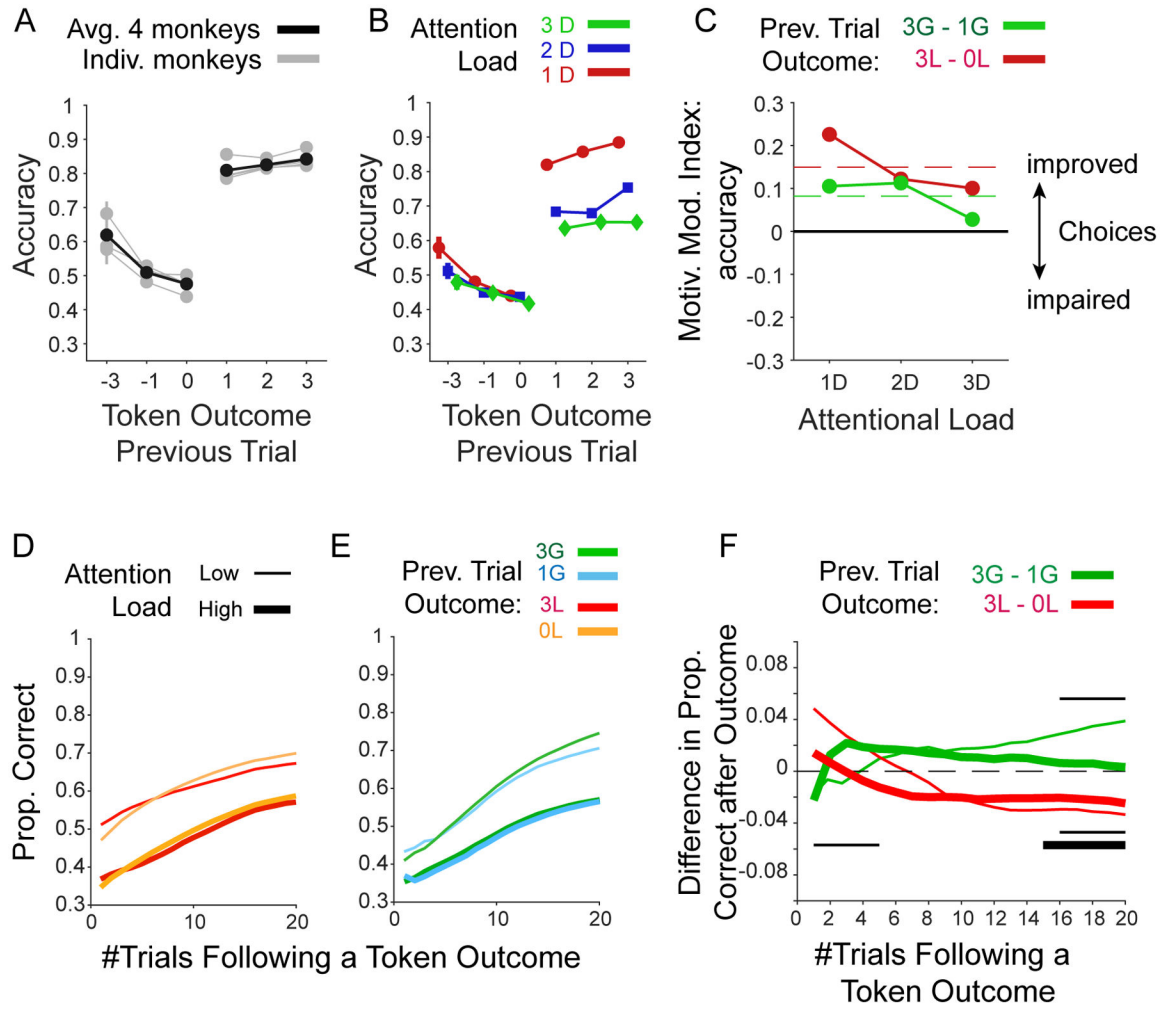
Figure 9. Effects of experienced token gains and losses on performance.
(A): The effects of an experienced loss of 3, 1 or 0 tokens and of an experienced gain of 1, 2, or 3 tokens (*x-axis*) on the subsequent trials' accuracy (*y-axis*). Grey lines are from individual monkeys, and black shows their average. **(B)** Same format as *A* for the average previous trial outcome effects for low/medium/high attentional load. **(C)** Motivational modulation index (y-axis) quantifies the improved accuracy after experiencing 3 vs. 1 losses (*red*), and after experiencing 3 vs. 1 token gains (*green*) for low/medium/high attentional load. Dashed red lines show grand mean across attention load conditions. **(D)** The effect of an experienced loss of −3 or 0 token during learning on the proportion of correct choices in the following n[th] trials (x-axis). **(E)** Same as *D* but for an experienced gain of 3 or 1 token. **(F)** The difference in accuracy (*y-axis*) after experiencing 3 vs. 0 losses (*red*), and 3 vs. 1 token gains (*green*) over n trials subsequent to the outcome (x-axis). Thick and thin lines denote low and high attentional load conditions, respectively. Black thin and thick horizontal lines in the lower/upper half of the panel show trials for which the difference of accuracy (high vs. low losses/gains) was significantly different than zero, Wilcoxon test,

FDR corrected for dependent samples across all trial points and low and high attentional load conditions with an alpha level of 0.05.

Table 1.

Model comparisons for different control conditions. The last row in each table is the model with gain and/or loss feedback, compared with the other models. BIC: Bayesian Information Criterion, AIC: Akaike Information Criterion. LR-Stat: Likelihood ratio stat.

Models on all conditions				
Model	**BIC**	**AIC**	**LR-Stat**	**P-value**
$RR_{stimulation}$	21210	21174	−49.69	<0.001
$RR_{Experienced}$	21203	21167	−42.35	<0.001
$Diff_{Gain-Loss}$	21165	21130	−58.84	<0.001
$FB_{Gain} FB_{Loss}$	21146	21081		
Models on fixed gain: 2, variable loss: 0,−1,−3				
Model	**BIC**	**AIC**	**LR-Stat**	**P-value**
RR	7462	7433	−58.20	<0.001
FB_{Loss}	7404	7375		
Models on fixed loss: −1, variable gain: 2,3				
Model	**BIC**	**AIC**	**LR-Stat**	**P-value**
RR	5723	5695	−1.04	0.14
FB_{Gain}	5722	5694		
Models on fixed gain: 3, variable loss: 0,−1				
Model	**BIC**	**AIC**	**LR-Stat**	**P-value**
RR	10384	10353	−7.29	<0.001
FB_{Loss}	10377	10346		
Models on fixed loss: 0, variable gain: 1,2,3				
Model	**BIC**	**AIC**	**LR-Stat**	**P-value**
RR	15356	15323	−3380	<0.001
FB_{Gain}	11975	11943		