Gains and Losses Differentially Regulate Attentional Efficacy and Learning 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.*

Corresponding Author: Dr. Thilo Womelsdorf (thilo.womelsdorf@vanderbilt.edu)

Keywords:
Feature-based attention, incentive salience, loss aversion, approach and avoidance, cognitive control, reinforcement learning

Abstract

Visual attention involves both enhancing the processing of stimuli that may lead to reward and avoiding the processing of stimuli that may lead to loss. But compared to reward expectancy little is known how loss-avoidance mediates attention. One hypothesis holds that attention is more efficiently deployed when expecting larger gains and larger losses as both increase motivational saliency. Alternatively, expecting larger penalties might reduce attentional efficacy and 'loom' larger than gains. Here, we tested these opposing views in four monkeys with a feature-token learning task that quantifies attentional efficacy by increasing distractor load from multidimensional objects. During learning the number of token rewards gained for correct choices and lost for incorrect choices were varied. We found that expecting larger gains improves attentional efficacy and learning speed as long as distractor load was low. In contrast, expecting larger losses impairs attentional efficacy and this impairment increases with distractor load and uncertainty about the relevance of visual

features. These findings functionally dissociate the contributions of expecting gains and losses on attentional learning, suggesting that they operate via separate control pathways. One pathway is linked to avoiding loss by slowing down learning and disrupting attention non-selectively, while the second pathway enhances learning selectively for higher valued target features up to a limited distractor load. These results illustrate the strength and the limitation of motivational regulation of attentional efficacy during learning.

Significance statement

Visual attention is biased to objects promising higher expected gains, but it should likewise be biased to avoid processing objects associated with incurring a loss. Clarifying the roles of gains and losses to modulate attention is critical for understanding the possible neuronal processes that determine how efficient attention is deployed in real-world learning scenarios. This study clarifies that gains and losses have surprisingly strong, opposite effects on the efficacy of attentional deployment in environments with low and high distractor load. Attentional efficacy is enhanced with increased gains primarily when there are few distractors, while attentional efficacy is impaired with increased loss expectancy particularly when there are many distractors. These results clarify that avoiding loss and striving for gains affect attentional efficacy via partly distinct mechanisms.

Introduction

Attention prioritizes the processing of sensory information though mechanisms that enhances relevant and suppresses non-relevant information (Womelsdorf and Everling, 2015; Noonan et al., 2018). In visual search and related tasks, such mechanisms include motivational factors that

enhance the representation of target features when they are linked to positive reward expectations (Bourgeois et al., 2016; Anderson, 2019; Theeuwes, 2019), and incentive salience (Berridge and Robinson, 2016; Failing and Theeuwes, 2018). While such incentive-related enhancements are broadly accepted, the effects of losses on attention are less well-understood. In contrast to target enhancement through incentives, loss related mechanisms are considered essential to avoid aversive outcomes by suppressing distracting visual information (Noonan et al., 2018; Chelazzi et al., 2019). When learning about the relevance of visual objects, neutral objects become distractors through mere over-familiarization, but pairing objects with unpleasant consequences (such as a bitter taste or a monetary loss) can induce avoiding to look at these stimuli, which presumably reflects avoiding to attend to these stimuli (Raymond and O'Brien, 2009; Ghazizadeh et al., 2016), rendering aversive- or loss- associated distractor objects easier to ignore in some situations (Raymond and O'Brien, 2009; Gupta et al., 2016), but not all situations (Becker et al., 2020).

The observation of active suppression of specific visual objects with aversive associations suggests that the de-prioritization of distractors acts in parallel with the active prioritization of target representations during visual attention (Noonan et al., 2018). Our study set out to test this assertion by quantifying how the experience of both, gains and losses modulates how efficiently visual distractor features are suppressed before and after learning. There are various conflicting hypotheses about how losses and gains should influence distractor interference in such a situation. One hypothesis suggests that experiencing loss should be as effective as experiencing gains in orienting attention to and learning about a stimulus as long as losses and gains are of similar magnitude. According to this idea it is the surprise, or absolute unexpectedness of outcomes that determines how much attention is deployed and how much is learned about a stimulus (Pearce and

Hall, 1980; Esber and Haselgrove, 2011). Alternative views hold that attention is affected differently by gains and losses. In frameworks of avoidance learning, successful (instrumental) avoidance of a distracting aversive object is a positive outcome to subjects that trigger its own positive prediction error and updating process to more likely prevent processing that stimulus in the future (Maia, 2010). According to this framework, expecting loss should trigger active learning to avoid aversive distractors which should improve the control of distractor interference. An opposite prediction is made by rational theories of cognitive effort that assert allocating attention to stimuli depend on the ratio of the expected benefits (gains) over the expected costs (losses) associated with stimuli (Shenhav et al., 2013). Since aversive outcomes represent a cost it should reduce attentional control and make subjects more vulnerable to distractor interference, while expecting gains should increase attentional control (Shenhav et al., 2017).

The outlined scenarios do not agree on how motivational factors affect the efficient suppression of distraction during attentive processing. Gains and Losses might have opposite effects on attention and distractor suppression (Shenhav et al., 2017), or alternatively averting loss might be as effective as positive reward expectations to enhance attention and mitigate interference (Engelmann et al., 2009; Aarts et al., 2010; Krebs et al., 2010; Padmala and Pessoa, 2011; Umemoto and Holroyd, 2015). Here, we set out to test these views by measuring how expected gains and losses modulate the influence of increasing numbers of distractors on learning the identity of a rewarded target feature of multidimensional objects. The target feature was learned within blocks of 35-60 trials through trial-and-error by choosing one among three objects composed of multiple different object features (**Fig. 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). We achieved this parametric variation of distractibility using multidimensional Quaddle objects that varied features of 1, 2, or 3 dimensions from trial-to-trial (**Fig. 1C**)(Watson et al., 2019a). In addition to load, we orthogonally 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. condition with 3 gain and 1 loss tokens), while other conditions had lower reward rate despite the absence of loss tokens (e.g. condition with 1 gain and 0 loss tokens). This arrangement allowed studying losses relatively independent of overall reward rate (**Fig. 1D**,**E**). We found that expecting larger loss reduces the efficacy of learning targets most strongly at high attentional load, while expecting larger gains enhanced the efficacy of learning targets particularly at low attentional load.

Results

Four monkeys performed the feature learning task and collected tokens as secondary reinforcers to be cashed out with fluid reward when five tokens were collected. Monkeys tracked their earned tokens, as demonstrated by slower choice response times the more tokens were already earned during the time when the target feature was not yet learned (**Supplementary Fig. S1A**,**C**), and faster choice response times the more tokens they had earned after learning during asymptotic performance (**Supplementary Fig. S1B**,**D**). On average subjects completed 1080 learning blocks (SE ±32, range 1008-1166) in 30 test sessions (SE ±1, range 28-33). Each of the monkeys showed slower learning of the target feature with increased attentional load, and when expecting to lose more tokens for incorrect responses, but they increased the speed of learning the target feature the

more tokens they expected to earn for correct responses (**Supplementary Fig. S2 and S3**). These 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). In all linear mixed effects models monkeys (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 learning the target feature at varying distractor load and gains/losses as reflecting changes in the efficacy of attentional learning. Not only did monkeys learn faster at lower loads, when expecting less loss, and when expecting higher gains, but they also had fewer unlearned blocks under these same conditions (**Supplementary Fig. 4**).

The main effects of loss-expectancy and gain-expectancy provide apparent support for a valence-specific influence of motivational incentives and penalties on attentional efficacy (Shenhav et al., 2017). They are opposite to what is expected from saliency models that assume the absolute magnitude of outcomes determines attentional efficacy similarly (**Fig. 2A**). However, the valence-specific effect was not equally strong at low/medium/high attentional loads (**Fig. 2B**). We found that linear mixed effects models with interaction terms described the data better than without them (Likelihood Ratio Stat= 8.5, p=0.014, Bayesian Inform Crit. 21176 and 21169 for [$Att._{Load}$ x ($feedback_{Losses}$ + $feedback_{gains}$)], and [$Att._{Load}$ + $feedback_{Losses}$ + $feedback_{gains}$], 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 tokens for correct choices [MMI_{Gains} = $LEfficacy_{3G}$- $LEfficacy_{1G}$], and when expecting to lose 3 vs 0 tokens for incorrect choices [MMI_{Loss} = $LEfficacy_{3L}$- $LEfficacy_{0L}$]. By calculating the MMI

for each attentional load condition we visualize whether the motivation effect of increased gains and increased losses increases or decreases at higher attentional load (**Fig. 2C**). We found that the motivational effect from avoiding larger losses increased with attentional load, causing a larger decrease in efficacy at high load (**Fig. 2D**) (permutation test, $p < 0.05$). In contrast, the motivational effect from expecting higher gains improved learning most at low attentional load and had no measurable effect at high load (**Fig. 2D**) (permutation test, $p < 0.05$). Pairwise t-test comparison also showed MMI_{Loss} was significantly different from zero (p=0.002, p=0.008, and p=0.001 for low, medium, and high load), while MMI_{Gains} was only significantly different from zero in low load gain conditions (p=0.007, p=0.087, and p=0.95 for low, medium, and high load).

These contrasting effects of gains and losses on learning efficacy were partly paralleled in post-learning accuracy. Accuracy was enhanced with higher expected gains at lower but not at the highest attentional load conditions (t-test pairwise comparison, p<0.001, p=0.01, and p=0.78 for low, medium, and high load), decreased with higher expected losses at all loads (t-test pairwise comparison, p=0.003, p=0.017, and p=0.001 for low, medium, and high load), but this decrease was not modulated by load level (permutation test, p>0.05) (**Fig. 3A**,**B**). In contrast to learning speed and post-learning accuracy, response times varied more symmetrically across load conditions. At low load choice times were fastest with larger expected gains, and slowest with larger expected token-loss (**Fig. 3C**,**D**). At medium and higher attentional loads these effects were less pronounced.

To understand how gains and losses modulated learning efficacy on a trial-by-trial level we calculated the choice accuracy in the trials immediately after experiencing a loss of 3, 1, or 0

tokens, and after experiencing a gain of 1, 2, or 3 tokens. Experiencing more negative outcomes caused all monkeys to improve performance in the trials immediately following the negative outcome (**Fig. 4A**). This improvement was particularly apparent in the low attentional load condition, and less in the medium and high attentional load conditions (**Fig. 4B**) (LME model predicting the previous trial effect on accuracy; [$Att._{Load}$ x $feedback_{Losses}$], b=0.01, p=0.002). We quantified this effect by taking the difference in performance for loosing 3 vs. 0 tokens, which confirmed that there was least benefit from larger penalties in the high load condition, reflecting reduced learning efficacy (**Fig. 4C**). Similar to losses, experiencing larger gains improved the next trials accuracy strongest at low and medium attentional load (LME predicting previous trial effect on accuracy: for [$Att._{Load}$ x $feedback_{Gains}$], b=-0.03, p<0.001) (**Fig. 4B**,**C**). So far, the previous trial analysis considered the accuracy only in the immediate trial following lost or gained tokens. To discern longer term effects, we calculated accuracy with a n-trial running average window for 1-20 trials following the outcome. The analysis showed that immediately after losing 3 tokens, accuracy was increased compared with no losses, however this effect reversed over 2 trials in the high load conditions, and over 6 trials in the low load conditions (**Fig. 4D**,**F**). For the gain-conditions the results were different. A larger gain of 3 (vs. 1) token led to longer-lasting improved performance. This improvement was longer lasting in the low than the high attentional load condition (**Fig. 4E-F**). These results closely resemble the main effects of losses and gains on the learning efficacy (**Fig. 2**) 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 the expectancy of gains and loss had opposite effects on learning a relevant target feature at increasing attentional load. Expecting to lose tokens reduced learning speed (trials to criterion), impaired retention (post-learning accuracy), and increased choice response times, while expecting to gain 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 attentional efficacy at low and medium attentional load but had no beneficial effects at high load. These findings were paralleled on the trial-by-trial level by reduced behavioral adjustment following errors after losing 3 compared to 1 tokens, and by a lack of improving performance after correct choices of targets in the high load condition.

These findings document apparent asymmetries of gain- and loss- expectancy on attentional functions. The negative effect of symbolically losing a token and the positive effect of symbolically gaining a token were apparent in all four monkeys. The inter-subject consistency of the token and load effects suggests that our results delineate a fundamental relationship between the efficacy of visual selective attention and motivational influences from gains and losses.

Experiencing loss reduces the efficacy of visual attention. 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 (**Supplementary Fig. 1**). This findings is consistent with prior studies in rhesus monkeys showing that gaining visual tokens for correct performance and losing visual tokens for incorrect performance modulates performance in choice tasks (Seo and Lee, 2009; Rich and Wallis, 2017; Taswell et al., 2018). This sensitivity is essential to functionally

separate the influence of negative and positive expectancies 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 in our task (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, see **Fig. 2B**). This finding shows that experiencing losing 3 tokens does not enhance the processing, or the remembering, of the chosen distractors, but rather severely impairs using information from the erroneously chosen distractors to adjust behavior. This finding had not only a large effect size, but it was also unexpected for two reasons. First, it was unexpected because monkeys and humans can show large learning rates for negative outcomes when they are separately analyzed from learning rates for positive outcomes (Frank et al., 2004; Frank et al., 2007; Seo and Lee, 2009; Caze and van der Meer, 2013), indicating that negative outcomes can be useful for updating value expectations; And secondly, it was unexpected because monkeys successfully learn avoiding to look at —and presumably attend— objects that are paired with negative (primary) reinforcers (such as a bitter taste) (Ghazizadeh et al., 2016). According to these prior findings, monkeys should have been more effective in avoiding choosing the same distractor features when encountering them again. Instead, the token loss reduced their capability to avoid the wrongly chosen distractor features.

The main reason for this unexpected impaired learning to avoid distractor features is likely the high distractor density in our task. When tokens are lost following erroneously choosing an object with 1, 2, or 3 object features, the subjects appear unable to assign negative credit to a specific feature but may rather more diffusely assign negative credit to features that were available at the

time of the erroneous choice. Consistent with such an impaired credit assignment we found variable effects on post-error adjustment (**Fig. 4B**,**C**) and reduced performance over the longer time scale after experiencing loss (**Fig. 4F**) with the negative effects increasing with more distractors in the higher attentional load condition (**Fig. 2**). According to this interpretation, penalties like losing tokens are detrimental to performance when they do not inform straightforwardly what type of features should be avoided. Various lines of evidence support this view. Firstly, when penalties are more immediately informative about which objects should be avoided, e.g. in tasks with binary choice options, learning from negative outcomes can be more rapid than learning from positive outcomes (Averbeck, 2017). Secondly, inferring information from negative feedback seems to require effective fronto-striatal connectivity, because in humans the capability to benefit from negative outcomes only develops with age and correlates with the strength of fronto-striatal connectivity (Camara et al., 2009; van den Bos et al., 2012). This reasoning suggests that the monkeys might be limited in the effective use of negative feedback similar to younger human children. Thirdly, using aversive outcomes in a feature non-selective way might incur a survival advantage in an evolutionary sense because it promotes generalizing from specific aversive cues (e.g. encountering a specific predator) to a larger family of aversive events (all predator-like objects) which can enhance fast avoidance responses in future encounters (Krypotos et al., 2015). 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 distracting objects or object features when multiple, multidimensional objects define the learning environment.

Expecting gains enhances attentional efficacy but can't compensate for distractor overload. We found that expecting 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 (**Fig. 2**). On the one hand, this finding provides further quantitative support that larger incentives can improve learning efficacy and enhance attentional prioritization (Berridge and Robinson, 2016; Ghazizadeh et al., 2016; Walton and Bouret, 2019). In fact, learning was most efficient across all conditions when the monkeys expected three tokens for correctly choosing the object with the rewarded feature while other objects only varied in the same feature dimension (1D, low attentional load condition). However, this incentivizing of attention did not improve learning at high attentional load, i.e. when the distracting objects had each a total of three different feature that varied trial-by-trial (**Fig. 2**). 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 that defines the target feature did not vary from low to high attentional load. The only difference between these conditions was the higher distractor load, suggesting that the monkeys might have reached a limitation in distractor interference that they could not compensate by mobilizing additional control resources.

But what is the source of this limitation to control distractor interference? One possibility is that when attention demands increase monkeys are limited in allocating sufficient control or 'mental effort' to overcome distractor interference (Walton et al., 2003; Rudebeck et al., 2006; Walton et al., 2007; Parvizi et al., 2013; Hosking et al., 2014; Klein-Flugge et al., 2016; Shenhav et al., 2017). Thus, subjects might perform poorer at high attentional load partly because they are not allocating

sufficient control on the task. Rational theories of effort and control suggest that subjects allocate control as long as benefits outweigh the costs of exerting more control for a more demanding task and perform poorer at higher demand if the benefits (e.g. the reward rate for correct performance) are not increased concomitantly (Shenhav et al., 2013; Shenhav et al., 2017). The strength of such a rational theory of attentional control is that it does propose a specific, biologically plausible feature 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 demand is required with more interfering information. This view provides a parsimonious explanation for our findings at high 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 high attentional load condition. Our results therefore provide quantitative support for a rationale 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 correspond to an increase of distractor features from 8 (at medium attentional load) to 26 (at high attentional load). It thus seems that neural cognitive control circuitry in the rhesus monkey is able compensate for distractor interference whenever the number of distractors does not exceed a specific number of interfering features (which lies somewhere between 8 and 26 in our study). The precise number can be viewed as a limit for the allocation of selective attentional control using token-based incentivizing.

Neurobiology underlying the differential roles of gains and losses on regulating attentional efficacy. Our study provides behavioral evidence that expecting gains and avoiding losses regulate

attentional functions differentially. These differential effects are likely realized by partly segregated neural processing pathways underlying averting loss and negative experiences on the one hand and motivational incentives on the other hand. Both of these processing pathways proceed along the fronto-striato-amygdala axis (Berridge and Robinson, 2016; Hu, 2016; Averbeck and Costa, 2017). For learning to avoid aversive or distracting outcomes the anatomical connection from basal amygdala to central amygdala is critically necessary (Namburi et al., 2015) and lesions of amygdala reduces avoidance of stimuli with aversive outcomes (De Martino et al., 2010; Ilango et al., 2014). In the striatum lesions abolish selectively the ability to use reward incentives to optimize decisions (Schmidt et al., 2008). Intermingled with reward incentive coding neurons in the striatum are neurons of the so-called no-go pathway that learn to activate in response to aversive, undesired stimuli - including distractor events - in order to inhibit action or attention to them (Burke et al., 2017; Arcizet and Krauzlis, 2018; Banaie Boroujeni et al., 2020). This inhibitory striatal pathway receives input from medial prefrontal cortex where localized clusters of neurons mediate avoidance behavior to stimuli with negative outcome associations (Amemori and Graybiel, 2012; Amemori et al., 2020) and represent negative prediction errors that are specific to incorrectly attended visual object features (Oemisch et al., 2019). Anatomically intermingled to loss-associated processing in all of these areas are neural populations that mediate approach and attention to stimuli that promises reward outcomes, including in medial prefrontal cortex (Kaping et al., 2011; Vaidya and Fellows, 2015), amygdala (Averbeck, 2017; Rudebeck et al., 2017), and striatum (Taswell et al., 2018; Oemisch et al., 2019). In summary, the neurobiological evidence suggests that segregated neural populations and pathways within the fronto-striato-amygdala network mediate the motivational effects of gains and losses on attentional functions. The results presented here provide strong behavioral evidence that these pathways have a dissociated

influences on prefrontal or fronto-parietal top-down representations (Womelsdorf and Everling, 2015; Bisley and Mirpour, 2019). It will be important to understand in future studies how dopaminergic, noradrenergic and cholinergic modulation mediate these distinct processes that determine attentional efficacy (Collins and Frank, 2014; Berridge and Robinson, 2016; Hassani et al., 2017; Azimi et al., 2019; Sokol-Hessner and Rutledge, 2019).

Conclusion. Taken together, our results document the interplay of motivational variables and attentional efficacy during learning. We first showed that attentional efficacy is reduced when distractor load is increased despite the fact that the complexity of the feature-reward target rule did not change. This finding illustrates that attentional 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 attentional efficacy when there were distracting feature 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 the key process involved in cognitive effort (Shenhav et al., 2017). Moreover, the negative effect of distractor interference was exacerbated by expecting a penalty in the form of tokens for wrong choices. This effect shows that negative feedback did not help to avoid distraction during learning and thus did not provide a source for enhanced attentional control in uncertain environments.

Materials and Methods

Experimental Procedures. All animal and 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 substitutes 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 all 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., in preparation). 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 Quaddles that have a parametrically controllable feature space, varying along four dimensions, including the color, body shape, arm types and surface pattern (Watson et al., 2019a). Throughout training, one combination of features was never rewarded and hence termed 'neutral', which was the spherical, uniform, grey Quaddle with straight, blunt arms (**Fig. 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 animals performed the feature learning task for receiving tokens until they filled the token history bar with five tokens for a cash out of fluid reward (**Fig. 1A**). All animals effortlessly transitioned from immediate fluid reward delivery for correct choices to 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 et al., 2019b).

Task Paradigm. The task required monkeys to learn in blocks of 35-60 trials a target feature by choosing one among three objects, each composed of multiple features (**Fig. 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 the blue square touch was accepted, three stimuli appeared on the screen randomly on 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 distanced 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.

Each experimental session consisted of 36 learning blocks. Each learning blocks included 35-60 trials. Four monkeys (B, F, I, and S) performed the task and completed 33/1166, 30/1080, 30/1080, and 28/1008 sessions and blocks, respectively. We used a token-based rewarded 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 dimensions and features. Stimuli were multi-dimensional 2D viewed Quaddles which could present 1 to 3 features other than a neutral Quaddle (**Fig. 1C**) (Watson et al., 2019a). Four different dimensions/features were used in the task design (Shape, Color, Arm, and Pattern) and in each session of learning features of three dimensions were selected as potential target and distractor features for the learning blocks of that session. In each learning block only one feature corresponded to the correct target. All objects had the same number of non-neutral features. Attentional load was indexed as the number of non-neutral features presented on each object, varying from 1 to 3. The targeted feature was un-cued and monkeys were required to find the targeted feature through trial and error in each learning block. The task rule (target feature) remained the same throughout a learning block. After each block of learning the task rule changed and monkeys had to explore again to find out the new rule.

Each correct touch was followed by a yellow halo around the stimulus as visual feedback (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 (**Fig. 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 blinked 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.

Experimental design and conditions. 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 (**Fig. 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% accurate performance criterion from 5 to 25 trials (the so called 'learning point' in a block). This is the range of learning points showed for the different attentional load conditions in 90% of blocks during training. 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 that would be possible with the conditions that have lower reward rate. Since gaining 5 tokens

were 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.

Analysis of learning. The improvement of accuracy over successive trials relative to the begin of a learning block reflects the learning curve, which we computed with a forward-looking averaging window with a trial length of 10 over trials in each block (**Supplementary Fig. 2**). 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, i.e. learnt, >75% of blocks (B=67%, F=75%, I=83%, and S=74%, *see* **Supplementary Fig. 4**). We computed a *post-learning accuracy* as the proportion of correct choices in all trials after the learning criterion was reached in a block (**Fig. 3**).

Statistical analysis. We first constructed a linear mixed effect model (Pinherio and Bates, 1996) that tests 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 (1D, 2D, and 3D 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). 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 for others. This LME had the form

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

The model showed significant main effects of all three factors and random effects laid inside the 95% confidence interval (**Supplementary Fig. 3**). To test for the interplay of motivational variables and attentional load we extended the model with the interaction effects for Att_{Load} x Fb_{Gain} and Att_{Load} x $Fb_{Loss.}$ to:

$$LT/Accuracy = Att_{Load} + Fb_{Gains} + Fb_{Loss} + Att_{Load} \text{ x } Fb_{Loss} + Att_{Load} \text{ x } Fb_{Loss} + (1|Monkey) + (1|Feat) + b + \varepsilon \quad \text{(eq. 2).}$$

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 ($Diff_{Gain\text{-}Loss} = Fb_{Gain} - Fb_{Loss}$) as predictor. Secondly, we tested the role of Reward Rate (*RR*) as predictor. However, both these two models were inferior to the LME's formulated in eq. 2 and showed lower BIC. 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 the 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 (1D, 2D, and 3D objects), 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).

$$Accuracy/RT = Learn_{State} + Att_{Load} + Fb_{Gains} + Fb_{Loss} + (1|Monkey) + (1|Feat) + b + \varepsilon$$

(eq. 3).

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

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 and shown in **Fig. 4**.

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

Analyzing the motivational modulation of attentional load. To evaluate the influence of increasing gains and increasing losses on attentional load we calculated Motivational Modulation Indices (MMI) as the difference in learning speed (the trial-to-criterion) in the conditions with 3 versus 1 token gains (MMI_{Gain}) or in conditions with 3 versus 0 token losses (MMI_{Loss}). MMI_{Gain} and MMI_{Loss} were calculated separately for the low, medium and high attentional load condition. We statistically tested these indices under the null hypothesis that they are not different from zero by performing pairwise t-tests on each gain or loss condition given low, medium, and high attentional load (1D, 2D, and 3D).

To validate the t-test MMI results at individual load conditions, we used a permutation approach that tested the MMI across attentional load conditions. At the first step we randomly selected 100 samples from each feedback gain/loss condition and computed the MMI for each pair of samples separately for the feedback gain and for the feedback loss conditions in each attentional load. For

each attention load we selected 1000 random sub-samples with a sample size of 100 from each of the other 2 attentional loads separately, constructed 1000 95% confidence intervals and chose the top 5% confidence intervals to rectify the multiple comparison correction under the null hypothesis that MMI in each attentional load condition is not different from the each of other attentional load conditions.

Data and code availability

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.

Acknowledgments

The authors would like to thank Adam Neuman and Seyed-Alireza Hassani for help collecting the behavioral data. This research was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB028161 (T.W.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions

K.B.B. and T.W. conceived and performed the experiments. M.R.W. and K.B.B. wrote the code to run the experiments. K.B.B. analyzed the data. T.W., K.B.B., and M.R.W. wrote the paper.

Competing Interests

The authors declare no competing interests.

Figure Legends

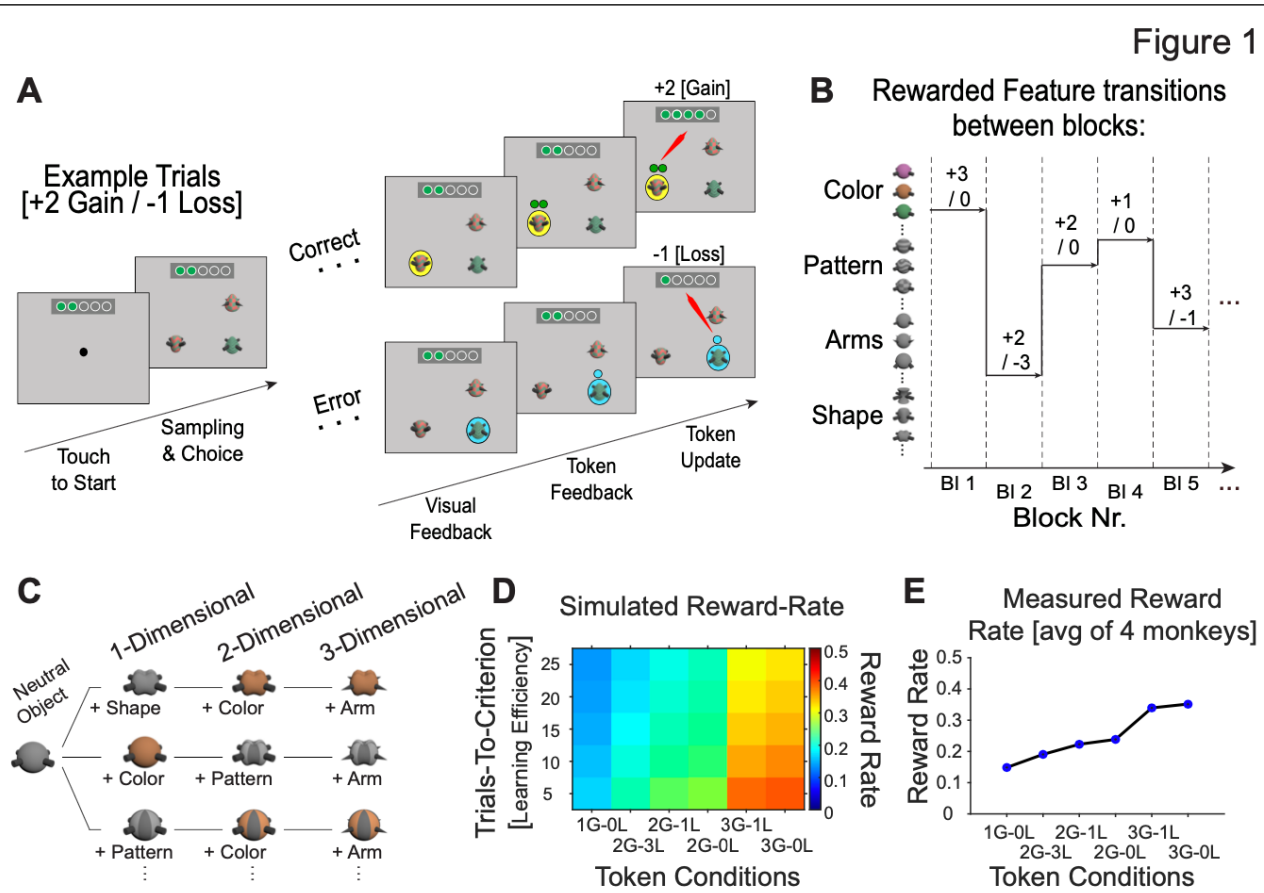


Figure 1. Task paradigm varying attentional load and token gains and losses. **A**: The trial sequence starts with presenting 3 objects. The monkey choses one by touching it. Following a correct choice a yellow halo provides visual feedback, then green tokens are presented above the stimulus for 0.3s 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.

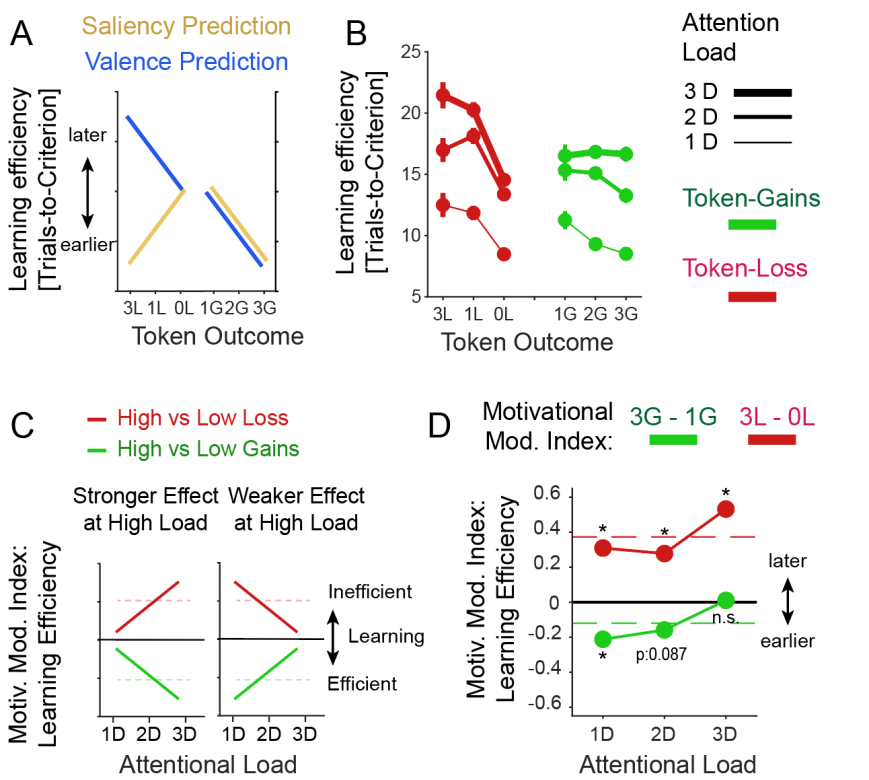


Figure 2. The effect of attentional load and expected token gain/loss on learning efficacy. **A:** The saliency hypothesis (*orange*) predicts that learning efficacy is high (learning criterion is reached early) when the absolute magnitude of expected loss (to the left) and expected gains (to the right) is high. The valence-specific hypothesis (*blue*) predicts that learning efficacy is high with high expected gains (to the right) and decreases systematically when larger penalties/losses are expected to the left. **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*).

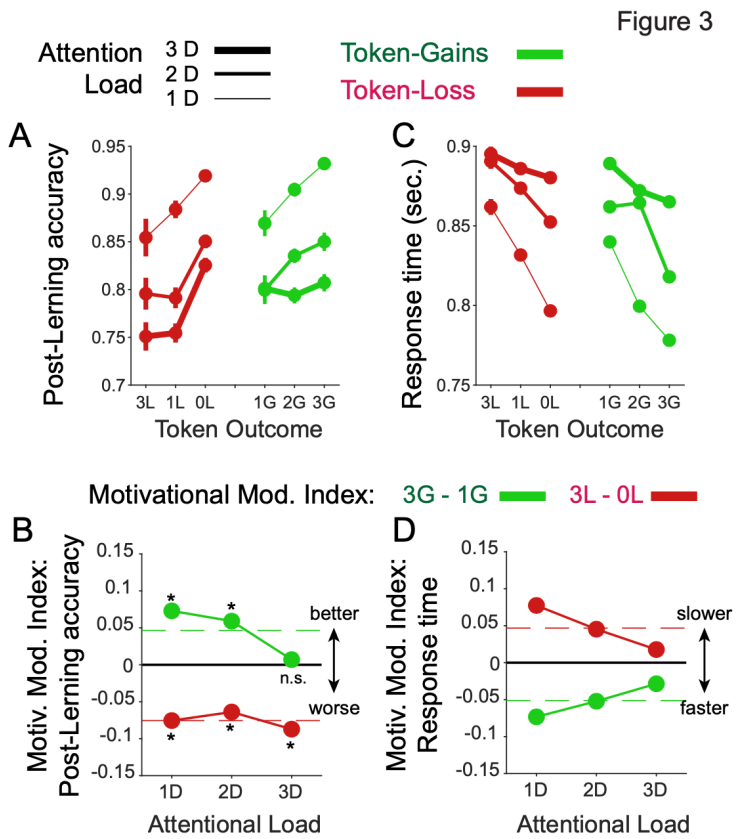


Figure 3. The effect of attentional load and token gain/loss expectancy on post-learning performance and response times. A: Post-learning accuracy (*y-axis*) when expecting varying token loss (*red*) and gains (*red*) at low/medium/high attentional load (line thickness). Overall, learning efficiency 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).

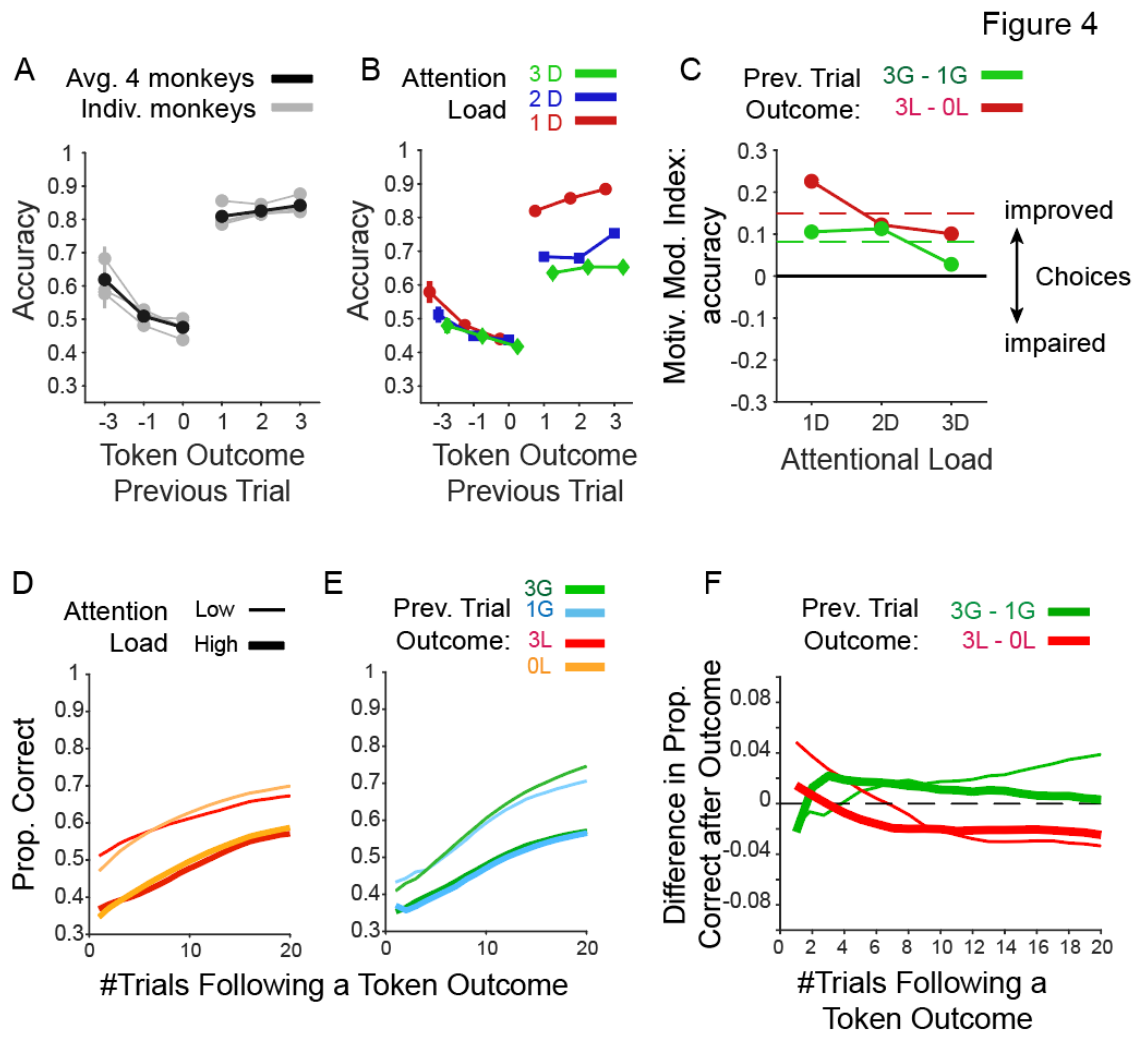
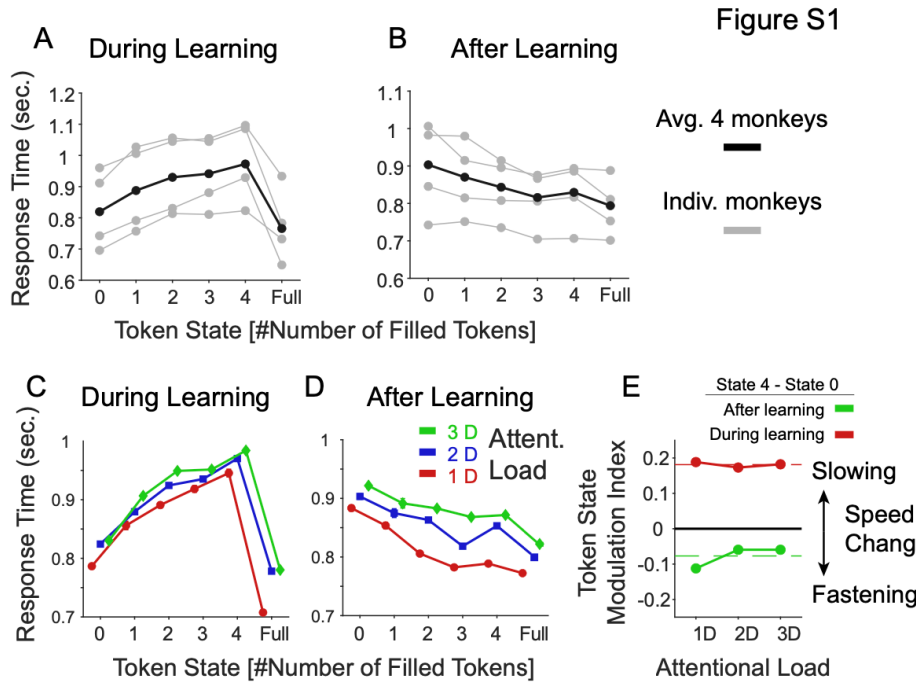
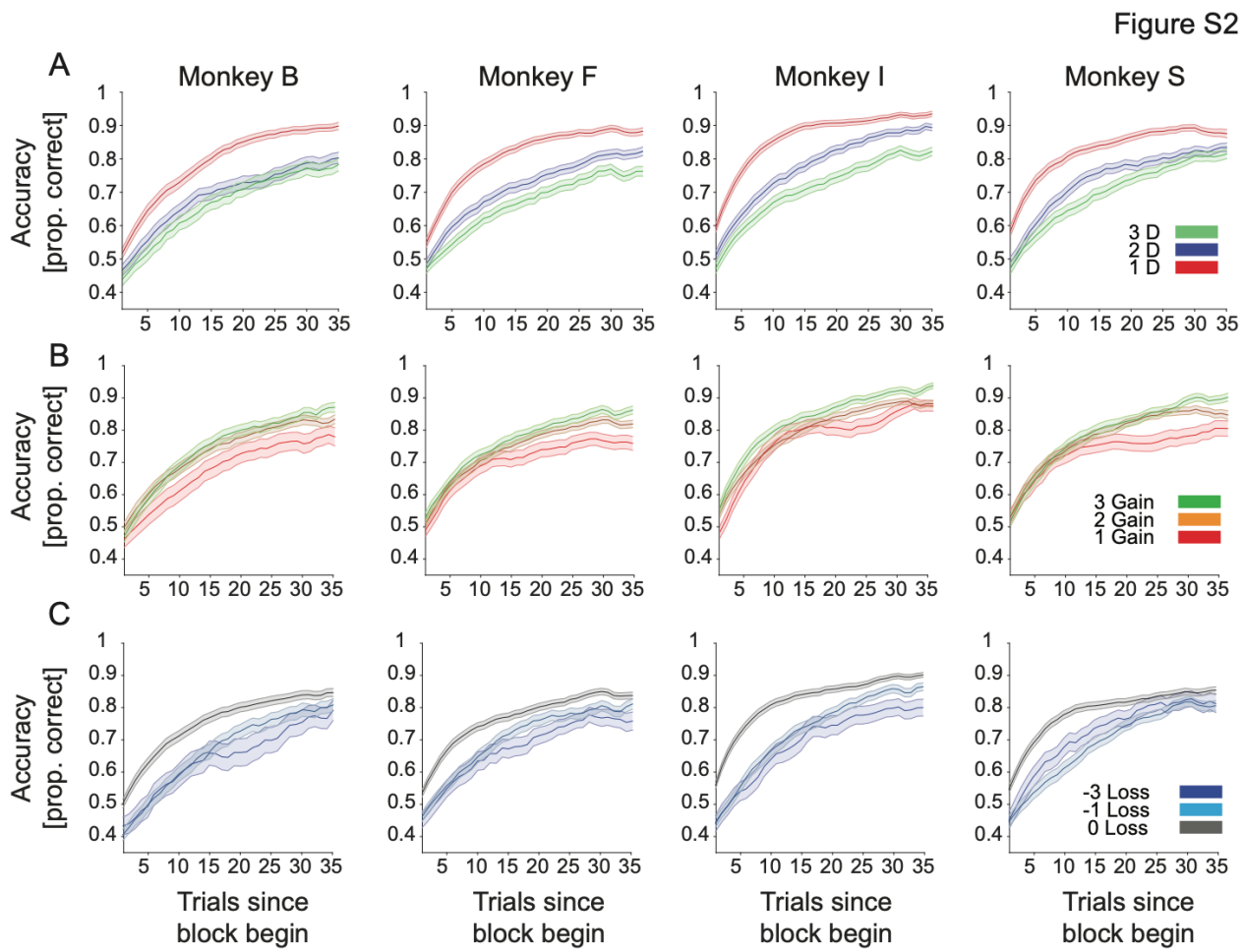


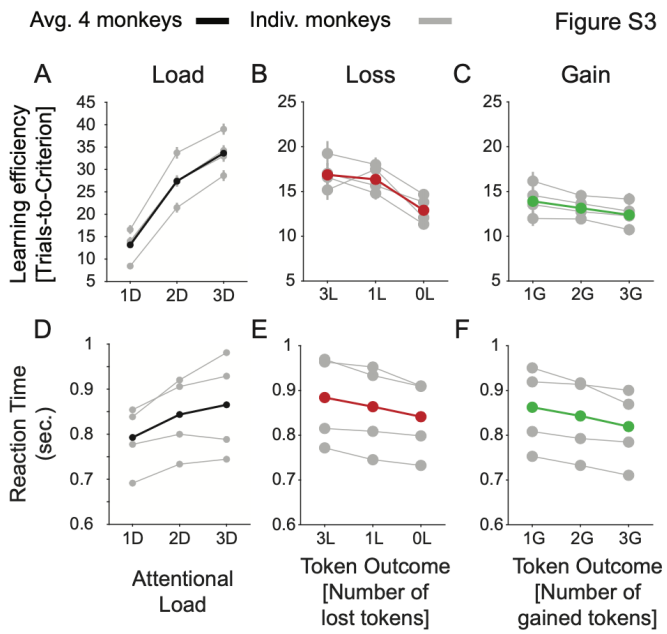
Figure 4. Immediate effects of experienced token gains and losses on performance. **A**: The effects of an experienced loss of 3, 1 or 0 tokens and of an experienced gain of 1, 2, or 3 tokens (*x-axis*) on the subsequent trials' accuracy (*y-axis*). Grey lines are from individual monkeys, and black shows 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. **D**: The effect of an experienced loss of -3 or 0 token 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.



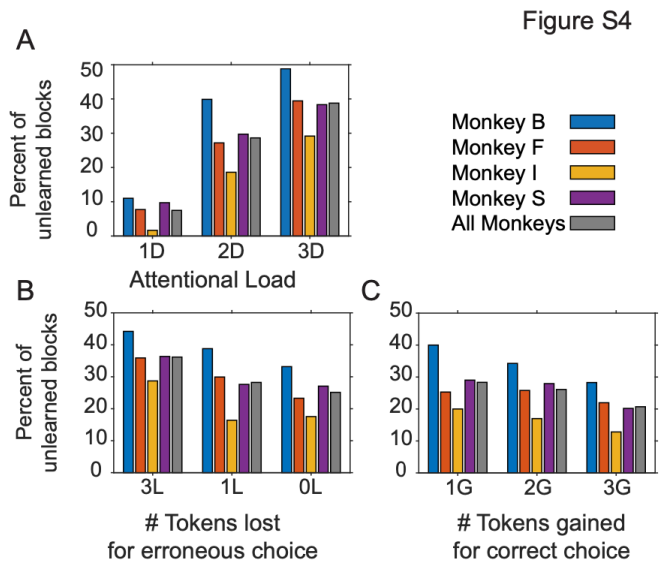
Supplementary Figure 1. Effects of the number of earned tokens ('Token State') on response times. **A**: Average response times as a function of the number of earned tokens visible in the token bar (*x-axis*) for individual monkeys (grey) and their average (in black). Included were only trials prior 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 (D). **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*).



Supplementary Figure 2. 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:** 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. **C:** 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.



Supplementary Figure 3. Main effects of attentional load, number of expected token-gains and expected token-loss on learning speed and response time. **A**: 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*). **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. **D-F:** Same as *A-C* for response times (*y-axis*).



Supplementary Figure 4. 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 conditions (*B*) and for the gains conditions (*C*).

References:

Aarts E, Roelofs A, Franke B, Rijpkema M, Fernandez G, Helmich RC, Cools R (2010) Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. Neuropsychopharmacology 35:1943-1951.

Amemori K, Graybiel AM (2012) Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. Nat Neurosci 15:776-785.

Amemori S, Amemori KI, Yoshida T, Papageorgiou GK, Xu R, Shimazu H, Desimone R, Graybiel AM (2020) Microstimulation of primate neocortex targeting striosomes induces negative decision-making. Eur J Neurosci 51:731-741.

Anderson BA (2019) Neurobiology of value-driven attention. Curr Opin Psychol 29:27-33.

Arcizet F, Krauzlis RJ (2018) Covert spatial selection in primate basal ganglia. PLoS Biol 16:e2005930.

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.

Averbeck BB, Costa VD (2017) Motivational neural circuits underlying reinforcement learning. Nat Neurosci 20:505-512.

Azimi M, Oemisch M, Womelsdorf T (2019) Dissociation of nicotinic alpha7 and alpha4/beta2 sub-receptor agonists for enhancing learning and attentional filtering in nonhuman primates. Psychopharmacology (Berl).

Banaie Boroujeni K, Oemisch M, Hassani SA, Womelsdorf T (2020) Fast spiking interneuron activity in primate striatum tracks learning of attention cues. Proc Natl Acad Sci U S A 117:18049-18058.

Becker MW, Hemsteger SH, Chantland E, Liu T (2020) Value-based attention capture: Differential effects of loss and gain contingencies. J Vis 20:4.

Berridge KC, Robinson TE (2016) Liking, wanting, and the incentive-sensitization theory of addiction. Am Psychol 71:670-679.

Bisley JW, Mirpour K (2019) The neural instantiation of a priority map. Curr Opin Psychol 29:108-112.

Bourgeois A, Chelazzi L, Vuilleumier P (2016) How motivation and reward learning modulate selective attention. Prog Brain Res 229:325-342.

Burke DA, Rotstein HG, Alvarez VA (2017) Striatal Local Circuitry: A New Framework for Lateral Inhibition. Neuron 96:267-284.

Camara E, Rodriguez-Fornells A, Ye Z, Munte TF (2009) Reward networks in the brain as captured by connectivity measures. Front Neurosci 3:350-362.

Caze RD, van der Meer MA (2013) Adaptive properties of differential learning rates for positive and negative outcomes. Biol Cybern 107:711-719.

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.

Collins AG, Frank MJ (2014) Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. Psychol Rev 121:337-366.

De Martino B, Camerer CF, Adolphs R (2010) Amygdala damage eliminates monetary loss aversion. Proc Natl Acad Sci U S A 107:3788-3792.

Engelmann JB, Damaraju E, Padmala S, Pessoa L (2009) Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects. Front Hum Neurosci 3:4.

Esber GR, Haselgrove M (2011) Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. Proc Biol Sci 278:2553-2561.

Failing M, Theeuwes J (2018) Selection history: How reward modulates selectivity of visual attention. Psychon Bull Rev 25:514-538.

Frank MJ, Seeberger LC, O'Reilly R C (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306:1940-1943.

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 104:16311-16316.

Ghazizadeh A, Griggs W, Hikosaka O (2016) Ecological Origins of Object Salience: Reward, Uncertainty, Aversiveness, and Novelty. Front Neurosci 10:378.

Gupta R, Hur YJ, Lavie N (2016) Distracted by pleasure: Effects of positive versus negative valence on emotional capture under load. Emotion 16:328-337.

Hassani SA, Oemisch M, Balcarras M, Westendorff S, Ardid S, van der Meer MA, Tiesinga P, Womelsdorf T (2017) A computational psychiatry approach identifies how alpha-2A noradrenergic agonist Guanfacine affects feature-based reinforcement learning in the macaque. Sci Rep 7:40606.

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 39:1558-1567.

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

Hu H (2016) Reward and Aversion. Annu Rev Neurosci 39:297-324.

Ilango A, Shumake J, Wetzel W, Ohl FW (2014) Contribution of emotional and motivational neurocircuitry to cue-signaled active avoidance learning. Front Behav Neurosci 8:372.

Kaping D, Vinck M, Hutchison RM, Everling S, Womelsdorf T (2011) Specific contributions of ventromedial, anterior cingulate, and lateral prefrontal cortex for attentional selection and stimulus valuation. PLoS Biol 9:e1001224.

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 36:10002-10015.

Krebs RM, Boehler CN, Woldorff MG (2010) The influence of reward associations on conflict processing in the Stroop task. Cognition 117:341-347.

Krypotos AM, Effting M, Kindt M, Beckers T (2015) Avoidance learning: a review of theoretical models and recent developments. Front Behav Neurosci 9:189.

Maia TV (2010) Two-factor theory, the actor-critic model, and conditioned avoidance. Learn Behav 38:50-67.

Namburi P, Beyeler A, Yorozu S, Calhoon GG, Halbert SA, Wichmann R, Holden SS, Mertens KL, Anahtar M, Felix-Ortiz AC, Wickersham IR, Gray JM, Tye KM (2015) A circuit mechanism for differentiating positive and negative associations. Nature 520:675-678.

Noonan MP, Crittenden BM, Jensen O, Stokes MG (2018) Selective inhibition of distracting input. Behav Brain Res 355:36-47.

Oemisch M, Westendorff S, Azimi M, Hassani SA, Ardid S, Tiesinga P, Womelsdorf T (2019) Feature-specific prediction errors and surprise across macaque fronto-striatal circuits. Nat Commun 10:176.
Padmala S, Pessoa L (2011) Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. J Cogn Neurosci 23:3419-3432.
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 80:1359-1367.
Pearce JM, Hall G (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol Rev 87:532-552.
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. Psychol Sci 20:981-988.
Rich EL, Wallis JD (2017) Spatiotemporal dynamics of information encoding revealed in orbitofrontal high-gamma. Nat Commun 8:1139.
Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF (2006) Separate neural pathways process different decision costs. Nat Neurosci 9:1161-1168.
Rudebeck PH, Ripple JA, Mitz AR, Averbeck BB, Murray EA (2017) Amygdala Contributions to Stimulus-Reward Encoding in the Macaque Medial and Orbital Frontal Cortex during Learning. J Neurosci 37:2186-2202.
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 7:327-336.
Schmidt L, d'Arc BF, Lafargue G, Galanaud D, Czernecki V, Grabli D, Schupbach M, Hartmann A, Levy R, Dubois B, Pessiglione M (2008) Disconnecting force from money: effects of basal ganglia damage on incentive motivation. Brain 131:1303-1310.
Seo H, Lee D (2009) Behavioral and neural changes after gains and losses of conditioned reinforcers. J Neurosci 29:3627-3641.
Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79:217-240.
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.
Sokol-Hessner P, Rutledge RB (2019) The psychological and neural basis of loss aversion. Current Directions in Psychological Science 28:20–27.
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:E12398-E12406.
Theeuwes J (2019) Goal-driven, stimulus-driven, and history-driven selection. Curr Opin Psychol 29:97-101.
Umemoto A, Holroyd CB (2015) Task-specific effects of reward on task switching. Psychol Res 79:698-707.
Vaidya AR, Fellows LK (2015) Ventromedial Frontal Cortex Is Critical for Guiding Attention to Reward-Predictive Visual Features in Humans. J Neurosci 35:12813-12823.
van den Bos W, Cohen MX, Kahnt T, Crone EA (2012) Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. Cereb Cortex 22:1247-1255.

Walton ME, Bouret S (2019) What Is the Relationship between Dopamine and Effort? Trends Neurosci 42:79-91.

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 23:6475-6479.

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.

Watson MR, Voloh B, Naghizadeh M, Womelsdorf T (2019a) Quaddles: A multidimensional 3-D object set with parametrically controlled and customizable features. Behav Res Methods 51:2522-2532.

Watson MR, Voloh B, Thomas C, Hasan A, Womelsdorf T (2019b) USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents. J Neurosci Meth:1-35.

Womelsdorf T, Everling S (2015) Long-Range Attention Networks: Circuit Motifs Underlying Endogenously Controlled Stimulus Selection. Trends Neurosci 38:682-700.

Womelsdorf T, Thomas C, Parker J, Neuman A, König S, Banaie Boroujeni K, Hassani AS, Watson M, Hoffman KL (in preparation) A Cage-based Monkey Kiosk for Robust Cognitive Profiling and Enrichment