Strategic self-control of arousal boosts sustained attention

J.W. de Gee[1,2], Z. Mridha[1,2], M Hudson[1,2], Y. Shi[1,2], H. Ramsaywak[1,2], S. Smith[1,2], N. Karediya[1,2], M. Thompson[1,2], K. Jaspe[1,2], H. Jiang[1,2], W. Zhang[1,2], M. J. McGinley[1,2,3]
[1]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA; [2]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA; [3]Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA.
Address correspondence to: jwdegee@gmail.com, matthew.mcginley@bcm.edu

Abstract

Spontaneous fluctuations in arousal influence neural and behavioral responses to sensory stimuli. However, it is not known if and how animals regulate their arousal to respond to fluctuating behavioral demands. Here, we show results from mice performing an auditory feature-based sustained attention task with intermittently shifting task utility. We use pupil size to infer arousal across a wide range of states in a large behavioral cohort. We find that mice stabilize their arousal state near an intermediate and optimal level when task utility is high. In contrast, during epochs of low task utility they sample a broader and suboptimal range of arousal states. Using diffusion modeling, we further show that multiple computational elements of attention-based decision-making improve during high task utility and that arousal influences some, but not all, of them. Specifically, arousal influences the likelihood and timescale of sensory evidence accumulation, but not its quality. In sum, the results establish a specific attentional signature of adaptively self-regulated arousal and provide an experimental framework for understanding arousal self-regulation in brain disorders such as attention-deficit/hyperactivity disorder.

Introduction

Autonomic arousal, controlled by the neuromodulatory centers of the reticular activating system, fluctuates on multiple time scales, shaping sensory processing and decision-making behavior in profound ways (Aston-Jones & Cohen, 2005; Harris & Thiele, 2011; Lee & Dan, 2012; McGinley, Vinck, et al., 2015). Fluctuations in pupil size at constant luminance track global arousal state (Joshi & Gold, 2020; McGinley, Vinck, et al., 2015) and the activity of the underlying neuromodulatory systems, including noradrenaline (Breton-Provencher & Sur, 2019; de Gee et al., 2017; Joshi et al., 2016; Murphy et al., 2014; Reimer et al., 2016; Varazzani et al., 2015) and acetylcholine (de Gee et al., 2017; Mridha et al., 2021; Reimer et al., 2016). Optimal behavioral and neural detection of simple sounds occurs at intermediate levels of pupil-linked arousal (McGinley, David, et al., 2015; Schriver et al., 2018), whereas large pupil size is associated with exploration and task disengagement (Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011; McGinley, David, et al., 2015) and small pupil size with drowsiness or sleep (McGinley, David, et al., 2015; Yüzgeç et al., 2018). This 'inverted-U' shape, or so-called Yerkes-Dodson relationship,

of the effects of pupil-indexed arousal on task exploitation, versus exploration or rest, encapsulates the 'adaptive gain' theory of noradrenergic tone (controlled by locus coeruleus) (Arnsten & Li, 2005; Aston-Jones & Cohen, 2005; Berridge & Waterhouse, 2003; Kane et al., 2017; Pfeffer et al., 2021; Usher et al., 1999; Yerkes & Dodson, 1908).

A central and unsolved question about the role of arousal in brain function is whether arousal fluctuations are purely spontaneous, perhaps driven by internal metabolic and/or memory consolidation demands (Squire et al., 2015), or whether humans and animals can strategically regulate their arousal state based on task demands (Aston-Jones & Cohen, 2005). For example, by the adaptive gain theory, when task utility (reward expectation) is high, one should regulate towards a moderate arousal state to facilitate optimal task engagement. In contrast, when utility is low, one should upregulate to a high arousal state to facilitate exploration of alternatives or downregulate to a low arousal state to rest and consolidate. In line with this proposed adaptability of arousal control, the locus coeruleus receives strong projections from frontal regions including the orbital frontal cortex and anterior cingulate cortex (Arnsten & Goldman-Rakic, 1984; Aston-Jones & Cohen, 2005; de Gee et al., 2017; Joshi & Gold, 2022; Porrino & Goldman-Rakic, 1982) and has more complex organization than previously assumed (Breton-Provencher & Sur, 2019; Totah et al., 2018). If and how organisms strategically regulate their arousal state is currently unknown, partly due to the prior difficultly of quantitatively monitoring arousal state in awake animals and humans.

Sustained and feature-based attention is an important cognitive function that can be effectively modeled in mice (Robbins, 2002; L. Wang & Krauzlis, 2018). Self-control of arousal importantly impacts attention. For example, dysregulated self-control of arousal is a prominent aspect of attention-deficit/hyperactivity disorder (ADHD) and other developmental disabilities and psychiatric disorders (de Lecea et al., 2012; Sander et al., 2015; Zhao et al., 2022). In particular, arousal is thought be important in turning up and down the intensity of attention, referred to as attentional effort (Ghosh & Maunsell, 2021; Sarter et al., 2006). In sustained attention tasks, as with most decisions, the accumulation of decision-relevant evidence is implemented in a distributed choice network of brain areas (Shadlen & Kiani, 2013; Steinmetz et al., 2019; van Vugt et al., 2018). In the canonical model of judgments about weak sensory signals in noise, sensory cortex encodes the noise-corrupted decision evidence, and downstream association and motor cortices accumulate this noisy sensory signal into a decision variable that determines the behavioral choice (Bogacz et al., 2006; Shadlen & Kiani, 2013; Siegel et al., 2011; X.-J. Wang, 2008).

We here sought to test if and how animals exert self-control on their arousal state to adapt attention-based decision-making performance to its utility. We developed a 'sustained attention value' task for head-fixed mice and report behavioral and pupillary signatures of strategic attentional effort allocation in a large cohort (N=88 mice; n=1983 total sessions). The task manipulates coherent acoustic motion in time-frequency space, analogous to random-dot motion used extensively in the visual system (Newsome et al., 1989). We apply simple behavioral analyses, as well as tailored accumulation-to-bound modeling, of the interacting effects of task-utility and pupil-indexed arousal. We find that during periods of high task utility mice exhibit

multiple signatures of heightened attentional effort, some of which are aided by stabilization of arousal closer to an optimal mid-sized level.

Results

A feature-based sustained attention task with nonstationary task utility

To study the role of arousal in adaptive utility-based regulation of sustained attention, we developed a quasi-continuous detection task for head-fixed mice (**Fig. 1**). Mice were trained to detect coherent time-frequency acoustic motion (called temporal coherence (Shamma et al., 2011)) embedded at unpredictable times in an ongoing random cloud of tones (**Fig. 1A**). The task requires sustained and attentive listening to achieve high detection performance, due to the high-order feature complexity and perceptual difficulty of the attended stimulus. Mice were motivated to elicit lick responses by food scheduling and administration of sugar-water reward upon correct licks (during the temporal coherence signal). To suppress excessive licking, mice received a 14-second timeout upon licking during the pre-signal tone cloud. We manipulated the utility of performing the task within each experimental session by shifting the reward size back and forth between high (12 ul) and low (2 ul) values in consecutive blocks (5-7 blocks per session) of 60 signal presentations. Hereafter, we refer to the task as the 'sustained-attention value' task. We simultaneously recorded pupil size and walking speed as measures of brain and behavioral state, respectively (**Fig 1B**). See Methods for details.

We defined four simple measures of behavioral performance to quantify distinct aspects of performance in the task: (i) Overall response rate, which is the number of responses per second of total noise or signal sound. Overall response rate captures the level of task engagement and, in this quasi-continuous go/no-go detection task, is analogous to decision bias (also called criterion) from signal detection theory (Green & Swets, 1966). (ii) Discriminatory response rate, which is the number of responses per second of signal sounds minus the number of responses per second of noise sounds. Discriminatory response rate captures how much more mice responded during signal vs noise stimuli and is therefore analogous to sensitivity (also called d') from signal detection theory. We interpret block-based shifts in discriminatory response rate as a readout of shifts in the intensity of attention (also called attentional effort; (Ghosh & Maunsell, 2021; Sarter et al., 2006)). (iii) Reward rate, which is the fraction of 'trials' (pairs of tone cloud alone and temporal coherence signal) that ended in a hit (a response during the signal, which was followed by a sugar-water reward). Reward rate is analogous to accuracy and is affected by both the overall and discriminatory response rates. (iv) Reaction time (RT) on hit trials, which is the delay from signal onset to correct response. Like reward rate, reaction time is affected by both the efficiency of sensory evidence accumulation and overall tendency to respond. See Methods for details.

By designing the sustained-attention value task to be demanding of attentional resources, we hypothesized that mice would increase their attentional effort during blocks of high reward, to exploit the high utility, and reduce attentional effort in low-reward blocks, to conserve cognitive resources and engage in other activities. Two features of the behavioral data supported that the task places high demand on attentional resources. The first of these was the pattern with which

they learned the task, which required three training phases (Methods). In phase 1, the signal was 6 dB louder than the noise, several free-reward trials (not contingent on responding) were included, and reward magnitude was constant (5 mL) throughout the session. During this phase, mice learned to respond (lick) to harvest rewards: the overall response rate increased from ~0 to ~0.6 responses/s within just three sessions (**Fig. 1C**). Discriminatory response rate also quickly increased across the first few sessions (**Fig. 1E**). In phase 2, we introduced the block-based shifts in reward magnitude and trained the mice until they reached a performance threshold (Methods). Overall response rate was ~0.3 responses/s at the end of phase 2 (**Fig. 1C**). Discriminatory response rate also gradually increased (**Fig. 1E**). In phase 3, the final version of the task, signal and noise sounds were of equal loudness and the small number of classical conditioning trials were omitted. By this time, mice were sufficiently experienced with the task to maintain a relatively high overall response rate (**Fig. 1D**). However, since they could no longer rely on the small loudness difference between signal and noise sounds, their discriminatory response rate dropped to approximately zero (**Fig. 1E**). Re-learning the – now purely feature-based – signal detection was, by design, perceptually difficult, as indicated by the shallow learning curves for discriminatory response rate and reward rate (**Fig. S1F,I**). Thus, mice rapidly learned the task structure, but slowly learned the signal acoustic structure and adaptive attentional effort allocation (**Fig. 1C-F; Fig. S1E-L**; see Methods).

The second feature of the behavioral data that demonstrates that the task places high demand on attentional resources was that the overall and discriminatory response rates declined across the session duration (~75 minutes/session; see **Fig. 2**), indicating that they could not sustain consistently high levels of performance. Eighty-eight mice completed training and performed a total of 1983 sessions of the final task phase.

Due to the quasi-continuous temporal structure of the task, it is theoretically possible that mice could learn to match their lick times to the temporal statistics of signal occurrence, rather than attend to the temporal coherence per se. To rule out such a timing-based strategy, we trained a cohort of mice (N=10; n=142 sessions) in a psychometric variant of the task in which signal coherence was varied randomly from trial to trial, in addition to the block-based utility manipulation. Signal coherence was degraded by manipulating the percentage of tones in each chord that moved coherently through time-frequency space (**Fig. S1M**; see Methods). This manipulation is analogous to reducing spatial motion coherence in the classic visual random-dot motion task (Newsome et al., 1989). As would be expected if the mice employed feature-based attention, rather than a timing-based strategy, both discriminatory response rate and reward rate lawfully varied with signal coherence (**Fig. S1N,O**). Thus, mice attended to temporal coherence in the tone cloud, which is a complex, high-order auditory feature.

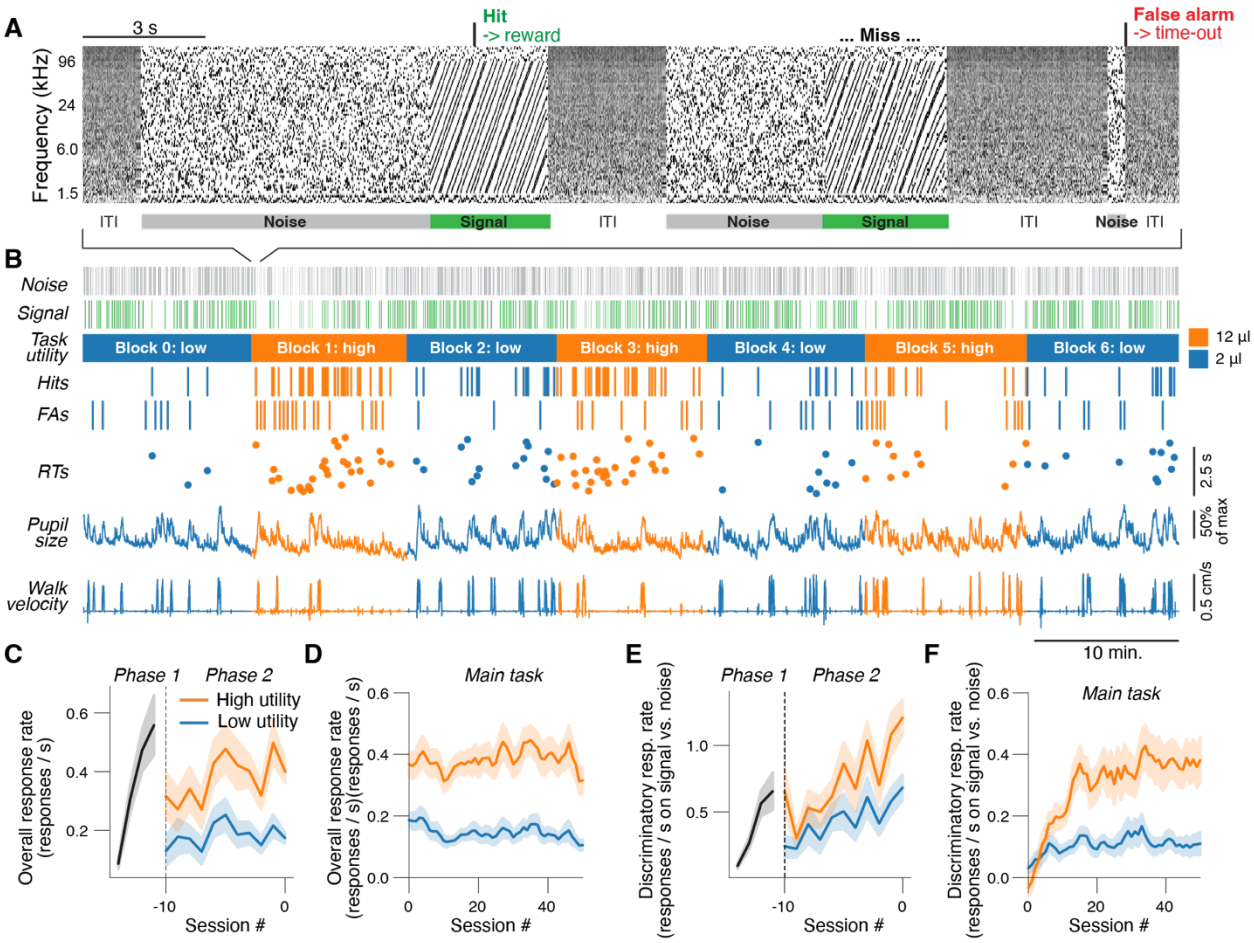

Figure 1. Monitoring performance and pupil-indexed arousal during the sustained-attention value task. (A) Spectrogram of the sound during an example of three consecutive trials. Correct go responses (hits, green text and arrow at top) were followed by 2 or 12 mL of sugar water. Reward magnitude alternated between 'low' and 'high' values (2 and 12 µL, respectively) in blocks of 60 trials. Incorrect go-responses (false alarms, red text and arrow at top) terminated the trial and were followed by a 14 s timeout, indicated to the animal as pink acoustic noise. **(B)** Example session. From top to bottom: noise stimulus timing, signal stimulus timing, reward/utility context, correct responses (hits), incorrect responses (false alarms), reaction times (RTs) on hit trials, pupil size and walking velocity. Color (blue or orange) indicates utility/reward context. **(C)** Overall response rate (Methods) across experimental sessions in learning phases 1 and 2 (Methods); session numbers are with respect to the last session in phase 2. **(D)** As C, but for the main task (Methods). See Fig. S1 for fitted learning curves and time constants. **(E,F)** as C,D but for discriminatory response rate.

Increase in task utility elicits rapid upregulation of attentional effort

We analyzed if mice strategically regulate attentional effort. Specifically, we sought to determine the time scales (both within and across blocks) on which mice adapted their overall, and discriminatory, response rates to shifts in task utility (**Fig. 2**). We observed that mice spent most of the 1st block (termed block '0'; always low reward) becoming engaged in the task (**Fig.**

S2A-C; see also **Fig. S3B**). Therefore, we focused all analyses on the subsequent six blocks (termed blocks '1' through '6'), for which the correct-like triggered reward size alternated in blocks between 12 µL (high reward) and 2 µL (low reward). See Methods for details.

Across blocks, we found that the animal's overall response rate was higher in each of the high, compared to neighboring low, reward blocks within an experimental session (**Fig. 2A-C**). Overall response rate was close to the optimal level, as determined by simulations (**Fig. S2E**; Methods), in the high reward blocks, while it was substantially lower than optimal in the low reward blocks (**Fig. 2C**). This pattern represents a performance, but not attention, optimization, because the overall response rate is not related to sensory signal detection, but rather reflects the animal's overall tendency to elicit responses, which is typically referred to as bias or criterion. However, discriminatory response rate was also higher in the high compared to low reward blocks, especially early in each session (**Fig. 2D-F**), indicating that feature-based attentional effort adapted to task utility. Both overall and discriminatory response rates declined across the session (**Fig. 2A,C,D,F**), probably resulting from effects of fatigue and/or satiety, both of which both slowly decrease the relative utility of performing the task as each session progresses. In sum, an increase in task utility boosts feature-based attentional effort, as indicated by discriminatory response rate, which is analogous to sensitivity. High task utility also boosts the signal-independent response probability (overall response rate, analogous to criterion) to a near-optimal level.

To understand how quickly mice could deploy attentional effort, we next sought to determine the within-block time course of these block-based behavioral changes. Because the reward context was not cued, the first reward in a block provided a strong and unambiguous indication that a transition in utility had occurred. Therefore, we analyzed time-dependencies within each block in two different ways: (i) aligned to the (unobserved) switch in reward availability (dashed lines in **Fig. 2A,D,G**), and (ii) aligned to the first hit trial after block transitions (solid lines). After the first large reward in the high reward block, overall response rate immediately increased by 232 ± 8%, compared to just before the switch, and did not further increase during the block (**Fig. 2B**). In contrast, when switching from high to low reward, overall response rate immediately decreased by 41 ± 2% and then decreased by a further 40 ± 2% with a time constant of 15 ± 1 trials (see Methods). For discriminatory response rate, after the first high reward in the high reward block, the rate immediately increased by 172 ± 18% compared to just before the switch, and then increased by a further 4 ± 4% with a time constant of 4 ± 2 trials (**Fig. 2E**). When switching from high to low rewards, discriminatory response rate immediately increased to a small extent, by 16 ± 7% (likely due to residual high attention from the previous high reward block, but decrease post-reward false alarms), and then decreased by 59 ± 2% with a time constant of 15 ± 1 trials. Thus, we found evidence for a hysteresis effect, similar to what has previously been observed in monkeys (Ghosh & Maunsell, 2021). Specifically, mice updated their behavior faster when switching from low to high reward, than the other way around, particularly for discriminatory response rate. This hysteresis indicates a heightened urgency when task utility is detected to have increased.

An animal or human optimally performing a task should allocate cognitive resources in a way that maximizes the rate of returns. We thus wondered if mice collected more rewards after increases in task utility. This is not trivially so in the sustained-attention value task; unlike in standard go/no-go tasks, in our quasi-continuous task a noise stimulus always precedes each signal stimulus. Thus, an overall response rate that is too high leads to more false alarms than hits and lowers the reward rate (**Fig. S2E**). Mice did ignore the noise and detect the signal on a larger fraction of trials in the high vs. low reward blocks and thus collected more rewards (**Fig. 2G-I)**. An analysis of individual differences supported this pattern of strategic reward-rate optimization; reward rate was strongly predicted by discriminatory response rate, but not overall response rate, at the per-mouse level (**Fig. S2J**). Mice on average sustained a stable reward rate of 27.6 ± 0.6% across the entire session during the high reward blocks (odd numbered), while reward rate was lower in the low reward blocks from the start (block 2) and further declined across the session (blocks 4 and 6; **Fig. 2G,I**).

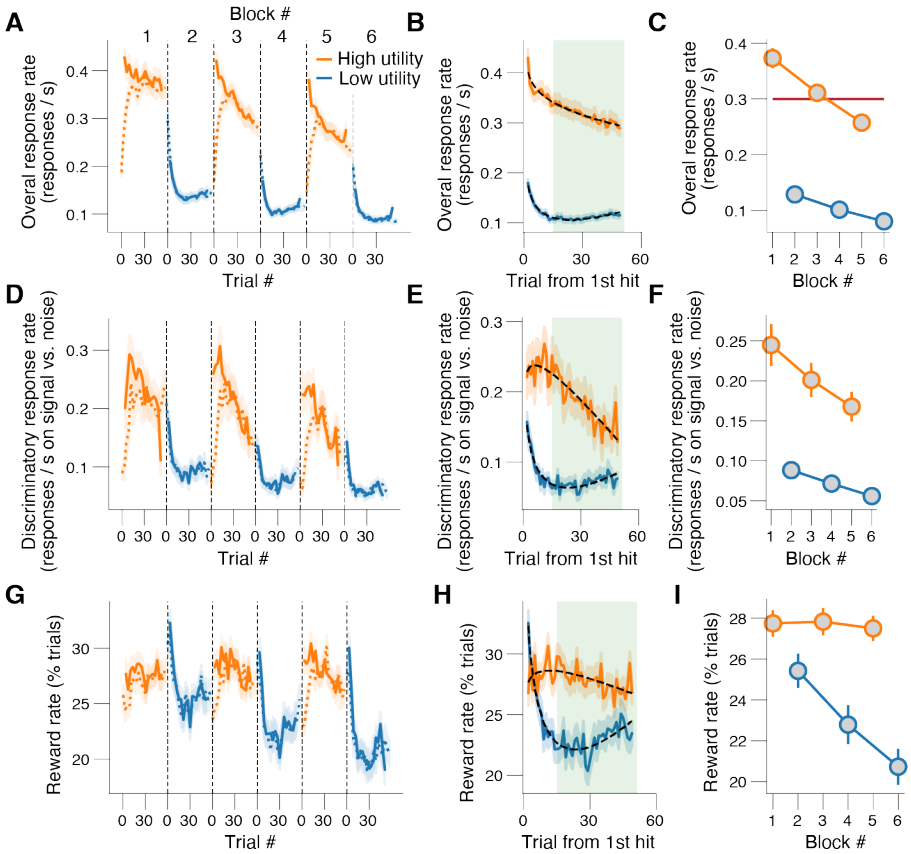

Figure 2. Rapid and adaptive changes in performance after shifts in task utility. (A) Overall response rate across three high-reward and three low-reward blocks in each experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(B)** As A, but collapsed across blocks of same reward magnitude. The green shaded area indicates the trials used when pooling data across trials within a block (e.g. panel C). **(C)** As A, but collapsed across trials within a block. Horizontal red line, optimal overall response rate (see Methods). Stats, 2-way repeated measures ANOVA (factors task utility [high vs. low] and time-on-task [1, 2, 3]); main

effect task utility: $F_{1,87} = 317.3$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 103.3$, $p < 0.001$; interaction effect: $F_{2,174} = 39.1$, $p < 0.001$. **(D-F)** As A-C, but for discriminatory response rate (Methods). Main effect task utility: $F_{1,87} = 63.0$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 17.1$, $p < 0.001$; interaction effect: $F_{2,174} = 4.0$, $p = 0.019$. **(G-I)** As A-C, but for reward rate (Methods). Main effect task utility: $F_{1,87} = 22.5$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 27.1$, $p < 0.001$; interaction effect: $F_{2,174} = 22.8$, $p < 0.001$. All panels: shading or error bars, 68% confidence interval across animals (N=88).

Mirroring reward rate, reaction times were substantially lower in high reward blocks than in low reward blocks, and gradually increased as the experimental session progressed (**Fig. S2F-I**). Furthermore, reward rate and reaction time exhibited similar hysteresis effects as the overall and discriminatory response rates (**Figs. 2H, S2H**). Specifically, mice updated all aspects of their behavior more quickly when switching from low to high reward than when switching in the other direction.

Further supporting that the attentional effort allocation was a perceptual decision-making, as opposed to procedural, optimization process, across animals the extent to which an animal allocated attentional effort was predicted by its overall discriminatory response rate, but not by its overall response rate **(Fig. S2K)**. We verified that our results were robust to specifics regarding trial selection (**Fig. S2L-O;** Methods) and were not confounded by effects of time-on-task (**Fig. S2P-S**; Methods). We conclude that mice increase their feature-based attentional effort after increases in task utility; when utility is high, they are more sensitive in discriminating signals from noise, are faster at doing so, and they collect more rewards.

Utility-driven attentional effort is partially mediated by stabilization of arousal

We previously showed that optimal signal-detection behavior in a simple tone-in-noise detection task occurred at intermediate levels of spontaneously fluctuating arousal (McGinley, David, et al., 2015). Based on these results, we sought to test whether the elevated attentional effort we observed during high task utility was afforded by the stabilization of arousal near its optimal state. To do so, we first identified the optimal arousal state in this task. We then determined whether mice spent more time close to this optimal state during periods of high task utility.

We quantified arousal as the diameter of the pupil (see **Fig. 1B**) measured immediately before each trial. We observed the highest discriminatory response rate and shortest RTs on trials with a mid-size pre-trial pupil size (**Fig. 3A,B**; see Methods for details). We additionally observed highest overall response rate on trials with mid-size pupil (**Fig. 3C**), and a linear, negative relationship between reward rate and pre-trial pupil size (**Fig. S3E**). The lack of inverted-U for reward rate results from the fact that the low overall response rate in the low pupil-linked arousal state resulted in fewer false alarms, and thus more trials on which the signal was played; this provided more opportunities to receive rewards. Locomotor status is another widely used marker of behavioral state (McGinley, David, et al., 2015; Polack et al., 2013). Overall response rate was

substantially higher on trials associated with pre-trial walking (**Fig. 3C**). However, discriminatory response rate was substantially lower and not statistically different from zero (asterisks in **Fig. 3A**). These results suggests that responses during walking were random (not signal-related), and thus that feature-based attentional effort was very low during locomotion.

We defined the optimal level of arousal as the pre-trial baseline pupil size bin with maximal discriminatory response rate (green vertical lines in **Fig. 3A-D**). Across animals, this optimal pre-trial baseline pupil size was 28 ± 2% of its maximum. We observed that during periods of high task utility mice spent more time close to this optimal arousal state (**Fig. 3D**). Strikingly, mice spent less time in both the low and high arousal states (blue and red shadings in **Fig. 3D**, respectively), both of which are suboptimal. To capture how close the animal's arousal state on each trial was to the optimal level, we computed the absolute difference between each pre-trial's pupil size and the optimal size (Methods). As the session progressed, their arousal state became less optimal (**Fig. S3F,H**), but, importantly, mice maintained their arousal closer the optimal state in the high-, compared to low-, reward blocks (**Fig. 3E** & **Fig. S3F,H)**. In line with this result, the pre-trial pupil size was overall more stable (less variable) in the high vs. low reward blocks (**Fig. S3Q-S**).

The raw pre-trial pupil size was also overall smaller in the high compared to low reward blocks (**Fig. 3F** & **Fig. S3I-K**). We tested whether changes in task utility were more associated with changes in the absolute difference between each pre-trial's pupil size and the optimal size (**Fig. 3E** & **Fig. S3F,H**) or with changes in the raw pre-trial pupil size (**Fig. 3F** & **Fig. S3I-K**). To do so, we first compared the effect sizes of the main effects of task utility on both arousal measures. The partial η^2 was 0.33 for pre-trial pupil size and 0.55 for its distance from optimal, indicating a larger effect size for distance from optimal compared to raw pre-trial size. Secondly, we performed a logistic regression of block-wise reward magnitude (indicated as 0 or) on either z-scored block-wise pre-trial pupil size or its distance from optimal. The fitted coefficients were negative in both cases, but significantly more so for the measure of distance from optimality (**Fig. 3G**). Thus, during heightened task utility, mice do not stereotypically downregulate their arousal state, but instead up- or down-regulate their arousal closer to its optimal level, depending on whether their arousal was too low or too high, respectively, in the first place.

Like pre-trial pupil size, walk probability was lower in the high, compared to low, reward blocks (**Fig. S3U-W**). Our findings that both pupil-linked and walk-related arousal were higher in the low reward blocks indicate that mice did not use the low reward blocks simply to rest, but instead to engage in alternative, aroused and perhaps exploratory, behaviors. This result is consistent with the adaptive gain theory of locus coeruleus (Aston-Jones & Cohen, 2005), which plays a major role in pupil control (Breton-Provencher & Sur, 2019; de Gee et al., 2017; Joshi et al., 2016; Reimer et al., 2016; Varazzani et al., 2015).

Having observed that epochs of high task utility are associated with both a more optimal arousal state and increased behavioral performance, we wondered to what extent the arousal optimization could explain the utility-related performance effects. To address this, we tested for statistical mediation (Baron & Kenny, 1986) of arousal in the apparent effect of task utility on the different performance metrics (**Fig. 3H**). We found that (the indirect path of) block-wise increases

in task utility predicting block-wise decreases in distance from the optimal arousal state, in turn driving block-wise increases in discriminatory response rate, partially mediate the apparent effect of task utility on discriminatory response rate (**Fig. 3I**). Changes in pupil-linked arousal had a similar mediation effect on the relationship between task utility and overall response, but to a significantly lesser extent (**Fig. 3I**). Similar results of the mediation analysis were observed for pre-trial pupil size, pre-trial pupil size standard deviation, or walking probability (**Fig. S3L,P,T,X**).

Taken together, we conclude that regulating pupil-linked arousal towards an optimal level partially implements the adaptive behavioral adjustments that match attentional effort to task utility.

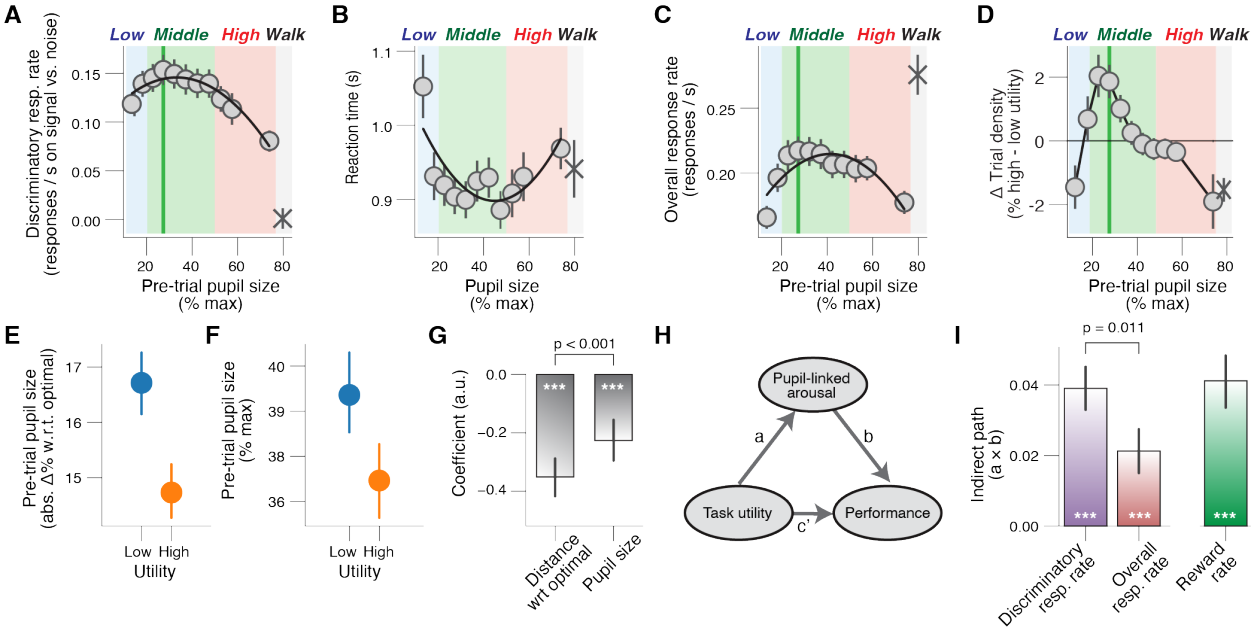

Figure 3. Mice regulate their arousal closer to optimality when task utility is high. (A) Relationship between pre-trial pupil size and discriminatory response rate (irrespective of task utility; Methods). A 1st order (linear) fit was superior to a constant fit ($F_{1,163} = 5.1$, $p = 0.024$) and a 2nd order (quadratic) fit was superior to the 1st order fit ($F_{1,963} = 4.0$, $p = 0.046$; higher order models were not superior; sequential polynomial regression; Methods). Grey line, fitted 2nd order polynomial; asterisk, walking trials (Methods); green vertical line, optimal pre-trial pupil size bin (see Methods). **(B)** As A, but for reaction time (RT) on hit trials. 1st order fit: $F_{1,163} = 2.5$, $p = 0.116$; 2nd order fit: $F_{1,963} = 4.9$, $p = 0.027$; higher order models were not superior. **(C)** As A, but for overall response rate (Methods). 1st order fit: $F_{1,163} < 0.1$, $p = 0.957$; 2nd order fit: $F_{1,963} = 8.3$, $p = 0.004$; higher order models were not superior. **(D)** Change in trial density after increases in task utility, separately for pupil-defined arousal states; asterisk, walking trials. **(E)** Absolute distance from optimal pre-trial pupil size separately for each block type. **(F)** As E, but for pre-trial baseline pupil size (Methods). **(G)** Coefficients from logistic regression of block-wise task utility [high vs. low] on absolute distance from optimal pre-trial pupil size or pre-trial pupil size. Stats, Wilcoxon signed-rank test; ***, p<0.001. **(H)** Schematic of mediation analysis of task utility to correct responses

(reward rate), via pre-trial pupil-linked arousal (Methods). Arrows, regressions; coefficient $a \times b$ quantifies the 'indirect' (mediation) effect; coefficient c' quantifies the 'direct effect'. **(I)** Fitted regression coefficients of the indirect path ($a \times b$; mediation). Stats, Wilcoxon signed-rank test; ***, $p < 0.001$. All panels: error bars, 68% confidence interval across animals (N=88, n=1983 sessions).

Multiple aspects of the attention-based decision computation improve when high task utility is high

We next sought to dissociate distinct elements of the decision-making process underlying the observed effects of task utility on overt behavior (see **Fig. 2**). Because the signal stimulus in our task was a high-order spectro-temporal statistic in ongoing noise, correct detection requires accumulation across time of partial evidence. Consistent with this perspective, reaction times were overall long and variable (**Fig. 4C,D**). We therefore applied sequential sampling modeling, as is commonly used for similar tasks in the primate visual system (Gold & Shadlen, 2007), or for auditory click discrimination in rodents (Brunton et al., 2013). It has been shown that rodents can perform acoustic evidence accumulation (Brunton et al., 2013; Hanks et al., 2015), but it is not known if and how evidence accumulation is shaped by task utility or arousal.

The widely used drift diffusion model describes the complete accumulation (i.e., without forgetting) of noisy sensory evidence as a decision variable that drifts to one of two decision bounds. Crossing a decision bound triggers a response, specifying a reaction time (Bogacz et al., 2006; Brody & Hanks, 2016; Laming, 1968; Ratcliff & McKoon, 2008). When the evidence is stationary (i.e. its summary statistics are constant across time), as occurs in typical go/no-go or two-alternative forced choice tasks, this model produces the fastest decisions for a fixed error rate (Bogacz et al., 2006). In our task, however, like perceptual decisions in most natural settings, the relevant evidence is not stationary. In this case, complete integration is suboptimal, because it results in an excessive number of false alarms due to integration of pre-signal noise (Ossmy et al., 2013). We thus used a computational model of the decision process based on leaky (i.e. forgetful) integration to a single decision bound (**Fig. 4A**; (Usher & McClelland, 2001); see Methods for details).

Our model contained five main parameters motivated by the design of the task and by general patterns observed in the behavior (**Fig. 4A**; Methods). The five parameters were: (i) leak, which controls the timescale of evidence accumulation; (ii) mean drift rate, which controls the efficiency of accumulation of the relevant sensory feature (temporal coherence in this case); (iii) attention lapse probability, which is the fraction of signal epochs for which the task-relevant sensory evidence is not accumulated; (iv) non-decision time, which is the speed of pre-decisional evidence encoding and post-decisional translation of choice into motor response; and (v) a mixture rate, which sets a fraction of trials on which the mouse makes an automatic response to noise (tone cloud) onset, and thus does not engage in evidence accumulation (**Fig. S4C,D**). Bound height was fixed to 1 (**Fig. 4A**; see Methods for details).

The attention lapse probability was included because correct responses (hits) were faster than expected for their frequency of occurrence. The mixture rate was included because we observed numerous very fast, but incorrect, responses immediately after tone cloud start. The fitted model accounted well for the behavior in the sustained-attention value task, making accurate predictions for both choice-outcome probabilities (**Fig. 4B**) and RTs (**Fig. 4C-D**). We considered four plausible alternative models (see Methods), including one with variable bound height, which provided worse fits, both qualitatively (**Fig. S4A-H**) as well as quantitatively (**Fig. S4I**).

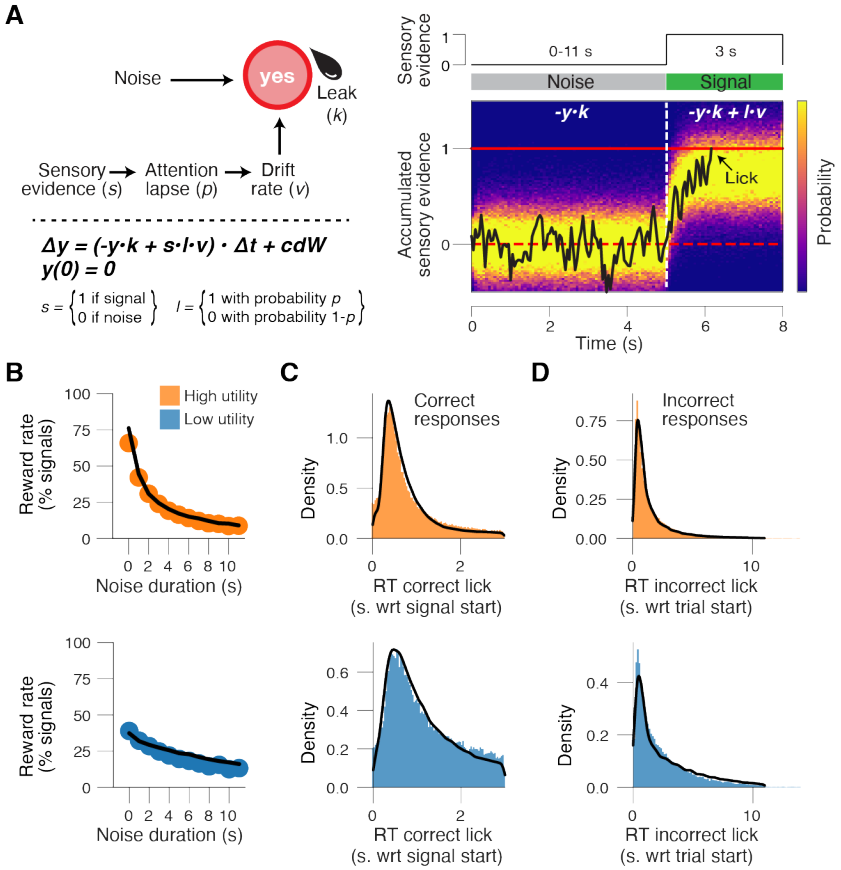

Figure 4. A leaky accumulation-to-bound model accounts for behavior in the sustained-attention value task. (A) Left: Schematic of bounded accumulation model accounting for the fraction of go-responses and their associated reaction times (RT). Right: The decision-dynamics were governed by leak (*k*) of gaussian noise during the tone cloud with attention lapse probability (*p*) and drift rate (*v*) during the signal stimuli. **(B)** Reward rate (defined as % correct responses during signal sounds) in the high (top) and low (bottom) reward blocks, separately for binned noise durations. Black line, model fit. **(C)** RT distribution for correct responses (hits) in the high (top) and low (bottom) reward blocks. Black line, model fit. **(D)** As C, but for incorrect responses (false alarms). Panels B-D: pooled data across animals (N=88) and sessions (n=1983).

Because there were prominent effects of both block-based task utility and of time-on-task in our overt behavioral measures (see **Fig. 2**), we first fitted the full model separately per block number (see Methods). We found that the drift rate was higher, and the leak and attention lapse probability were lower, in the high vs. low reward blocks (**Fig. 5A-C**). This pattern did not depend on the specifics of the model; for each of the alternative models, leak and attention lapse probability were lower in the high compared to low reward blocks (**Fig. S5E-G**). Drift rate was also higher in the high compared to low reward blocks when modeling the fast errors with variability in starting point instead of a mixture rate (**Fig. S5G**; Methods). There was no significant effect of task utility on the other, non-attention related, model parameters (**Fig. S5A,B**). The attention lapse rate increased, (**Fig. 5C**) and the mixture rate decreased (**Fig. S5B**), with time-on-task. In sum, increases in task utility resulted in a longer accumulation time constant (lower leak), more efficient evidence accumulation (higher drift rate) and more reliable evidence accumulation (lower attentional lapse probability).

Mid-level arousal has the lowest attentional lapse probability

We next fitted the full model separately for four arousal-defined bins: three pupil-defined bins during stillness, and a separate bin for walking trials (colored shadings in **Fig. 3A**; see Methods). Leak decreased monotonically with arousal (**Fig. 5D**), suggesting a longer timescale of evidence accumulation in higher arousal states. Attention lapse rate had a U-shaped dependence on arousal (**Fig. 5E**) indicating that the mice are most likely to engage with the signal (i.e. invest attentional effort) at moderate arousal levels. Additionally, we observed that automatic short-latency responses (mixture rate) increased monotonically with increasing arousal level (**Fig. S5D**). Thus, attention to the relevant feature is maximal at intermediate levels of pupil-linked arousal, and worst during walking.

The parameter dependencies of leak, drift rate, and attentional lapse rate on task utility all support the conclusion that attentional effort is greatest in high reward blocks and that arousal partially mediates the effect of reward, primarily by reducing attentional lapse rate. The higher drift rate and lower attention lapse probability in the high compared to low reward blocks contributed to a positive effect of task utility on discriminatory response rate and reward rate. Further supporting these conclusions, when stratifying the data based on animal-wise discriminatory response rates, we found substantial differences between strata in the same three attention-related parameters (**Fig. S5H-J**).

Lower leak indicates more sensory stimulus engagement, which is a signature of attentional effort. However, in our sustained-attention value task with nonstationary evidence, low leak does not necessarily increase reward rate, because more noise is accumulated. Therefore, we used simulations to identify the optimal leak for the task (**Fig. S4J,K**). Leak estimates were lower than optimal in the high reward blocks, and higher than optimal in low reward blocks, indicating that mice overshoot their leak in both reward contexts, but in opposite directions. The high integration time constant (1 / leak) in the high reward blocks, and during high arousal, indicates that they were more engaged with the sound stimulus (i.e. paying more attention), but also meant that extensive pre-signal noise was integrated, leading to an excess of false alarms and a high overall response rate. In sum, we found that an increase in task utility and moderate

arousal resulted in robust changes of the decision computation, at the block-to-block and animal-to-animal levels, all of which support improved attention.

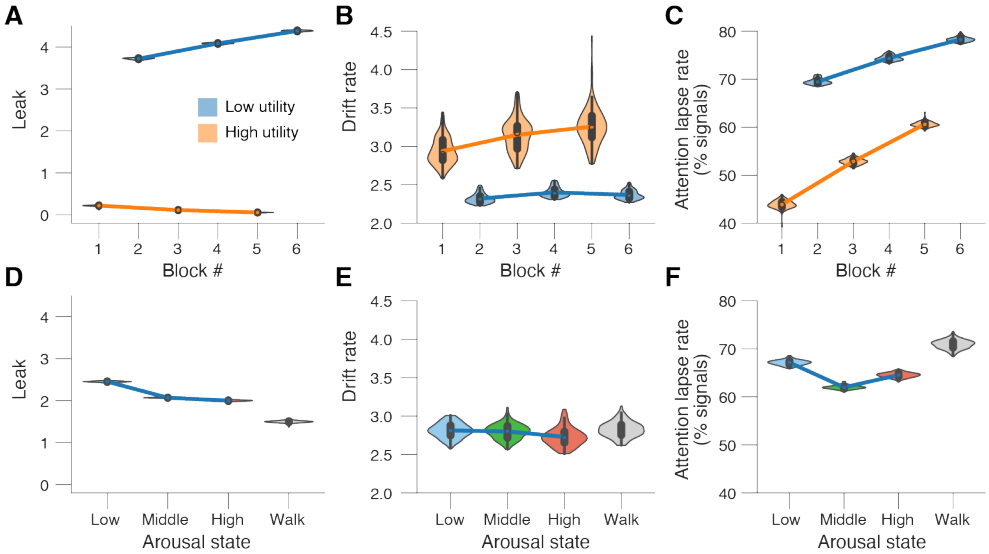

Figure 5. Task utility and arousal impact intersecting aspects of the decision computation. (A) Fitted leak estimates (kernel density estimate of 100 bootstrapped replicates) separately per block number. Main effect task utility (fraction of bootstrapped parameter estimates in the low reward blocks higher than in the high reward blocks): $p < 0.01$. Main effect time-on-task (fraction of bootstrapped parameter estimates in the first two blocks higher than in the last two blocks): $p = 0.50$. **(B)** As A, but for drift rate. Main effect task utility: $p < 0.01$. Main effect time-on-task: $p = 0.23$ **(C)** As A, but for attention lapse rate. Main effect task utility: $p < 0.01$. Main effect time-on-task: $p < 0.01$. **(D)** Fitted leak estimates (kernel density estimate of 100 bootstrapped replicates) separately per arousal state (pupil size defined bins corresponding to colors in **Fig. 3A-C**; irrespective of task utility; see Methods). Leak was significantly higher during low vs middle arousal ($p < 0.01$), during middle vs high arousal ($p < 0.01$) and during high arousal vs walking ($p < 0.01$). **(E)** As D, but for drift rate. No significant differences between arousal states. **(F)** As D, but for attentional lapse rate. The attentional lapse rate was significantly higher in the low vs middle arousal state ($p < 0.01$), and significantly lower during middle vs high arousal ($p < 0.01$) and during high arousal vs walking ($p < 0.01$). All panels: see fits of 'non-decision time' and 'mixture rate' in **Fig. S5A-D**.

Discussion

To efficiently meet their survival needs, organisms must regulate what Daniel Kahneman termed the *selective* and *intensive* (or effortful) aspects of attention (Kahneman, 1973). Characterizing the selective aspect of attention has been a cornerstone of systems neuroscience (Carrasco, 2011; Fritz et al., 2007; Maunsell & Treue, 2006), but attentional intensity/effort has received comparatively little scrutiny. Recent work emphasizes the importance of motivational factors in driving attentional effort (Brehm & Self, 1989; Ghosh & Maunsell, 2021; Kurzban et al., 2013; Richter et al., 2016; Sarter et al., 2006; Shenhav et al., 2013). This emphasis is in line with

the common experience of 'paying more attention' when motivated to perform better, for example in a classroom setting when the instructor indicates the next section will be on the final exam. In line with these ideas, heightened reward expectation increases perceptual sensitivity and reduces reaction times in humans and non-human primates (Engelmann & Pessoa, 2014; Ghosh & Maunsell, 2021; Locke & Braver, 2008). A rodent model of attentional effort has been needed so that underlying neural circuit mechanisms can be determined.

We here sought to meet this need, by developing a feature-based sustained attention task for head-fixed mice with nonstationary task utility. By applying quantitative modeling to a large behavioral and physiological dataset we show that during high utility, mice: (i) collect more rewards, (ii) accumulate perceptual evidence more efficiently, reliably, and across longer timescales, (iii) stabilize their pupil-linked arousal state closer to an optimal level; and (iv) suppress exploratory locomotor behavior. It has previously been shown that rodents can match their rate of learning to the statistics of a dynamic environment (Grossman et al., 2022), perform cost-benefit analysis (Reinagel, 2021), consider their own information processing limitations during strategic task performance (Grujic, Brus, et al., 2022) and adapt their response vigor to task utility (A. Y. Wang et al., 2013). However, it was not known if rodents can adapt attentional effort based on task utility. Our results add that regulating pupil-linked arousal towards an optimal level partially implements behavioral adjustments that adaptively increase attentional effort when it is useful.

Our results demonstrate an important, and previously underappreciated, role for motivation in driving attentional effort. Critically, mice sustained their highest level of performance – encapsulated in the reward rate – across the three high reward blocks interspersed across a long-lasting and difficult sustained attention task. Performance was comparably high in only the first low-reward block and then declined dramatically in subsequent low-reward blocks. A parsimonious interpretation of these findings is that early in the session mice are hungriest and least fatigued, and thus highly motivated to work for any reward. Later in the session, when more satiated and fatigued, only large rewards are sufficient to motivate them to increase their attentional effort (Hernández-Navarro et al., 2021). Another factor, which may increase performance in the low reward and blunt an even larger attention allocation effect, is that in our task design mice need to keep performing at a sufficiently high level to detect the switch from low to high reward block. Thus, the efficacy with which they can temporally allocate attention is probably higher than it appears in our results.

Growing evidence supports that neural computation and behavioral performance are not stationary within a session, due to ongoing fluctuations in internal state. State-dependence of neural activity has been observed in primary sensory cortices (Goris et al., 2014; McGinley, David, et al., 2015; Musall et al., 2019; Nestvogel & McCormick, 2022) and sensory-guided behavior (McGinley, David, et al., 2015) but not in relation to attentional effort or in any adaptive way. More recently, spontaneous shifts between engaged, biased, and/or disengaged states have been inferred from behavior (Ashwood et al., 2022; Weilnhammer et al., 2021). However, the behavioral functions of these non-stationarities is not unclear. In particular, it has not been demonstrated if and how animals adaptively control their arousal state to meet behavioral needs.

Utility-based shifts in behavior (but not attention) have been observed in rodents (Reinagel, 2021; A. Y. Wang et al., 2013). Other than in humans, the regulation of attentional effort by utility has only been observed in monkeys (Ghosh & Maunsell, 2021) and no study has assessed the role of arousal. Our results suggest that the large behavioral and neural variability that can be explained by fluctuations in arousal state serve an adaptive function; states conducive to a particular biological need (i.e. attention to a rewarded stimulus) are upregulated at appropriate times (i.e. when the reward is large). The adaptive function of low arousal states (such as for online learning and consolidation) (Pfeiffer, 2020; Squire et al., 2015) and high arousal states (such as for broadly sampling the environment to observe changes and exploring for alternatives) (Aston-Jones & Cohen, 2005), and self-regulation of sampling of these states, require further study.

An important question opened by our results is: 'which neuromodulatory systems contribute to the effects of arousal on attentional effort that we observed?' Our finding that tonic pupil-linked arousal is lower in the low reward blocks is in line with the adaptive gain theory of LC function (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011), but see (Bari et al., 2020), implicating norepinephrine. On the other hand, our results are not in line with the idea that increased acetylcholine mediates attention (Hasselmo & McGaughy, 2004; Sarter et al., 2006), but see (Robert et al., 2021). The willingness to exert behavioral control is thought to be mediated by tonic mesolimbic dopamine (Hamid et al., 2016; Niv et al., 2007; A. Y. Wang et al., 2013) and/or serotonin (Gutierrez-Castellanos et al., 2022). However, the willingness to work is likely more related to overall response rate, while attentional effort is more related to discriminatory response rate. Orexin/hypocretin is also involved in reward processing, and serotonin and orexin/hypocretin have both been implicated in pupil control (Cazettes et al., 2021; Grujic, Tesmer, et al., 2022). Future work is needed to determine the precise roles of neuromodulatory systems in utility-driven and arousal-regulated temporal allocation of attentional effort.

Our results demonstrating motivated shifts in attentional effort suggest a cost-benefit analysis being used to adapt arousal and performance level to an evolving motivational state (Botvinick & Braver, 2015; Shenhav et al., 2013). Neural underpinning of such top-down control of attentional effort are yet to be elucidated. Orbital frontal cortex (OFC) and dorsal anterior cingulate cortex (dACC) are likely candidate regions since both perform value computations related to optimizing behavior (Akam et al., 2021; Botvinick & Braver, 2015; Shenhav et al., 2013; Tremblay & Schultz, 2000). These structures are strongly connected to sensory cortices and to neuromodulatory nuclei, including LC (Arnsten & Goldman-Rakic, 1984; Aston-Jones & Cohen, 2005; de Gee et al., 2017; Joshi & Gold, 2022; Porrino & Goldman-Rakic, 1982). Impaired frontal regulation of arousal is one of the hallmarks of ADHD (Barkley, 1997) and plays a role in autism (Zhao et al., 2022) and a wide array of psychiatric disorders (de Lecea et al., 2012; Sander et al., 2015). Future work should determine the top-down influences of frontal cortices on arousal systems in attentional effort and their dysregulation in mouse models of neurological disorders.

Our observed role for pupil-indexed arousal in mediating adaptive adjustments in attentional effort is in contrast with the large literature on pupil dilation as a readout of attentional

capacity (also called effort) driven by fluctuations in task difficulty (Alnæs et al., 2014; Hess & Polt, 1964; Kahneman et al., 1967; Kahneman & Beatty, 1966; Laeng et al., 2012), including the extensive work on pupil-indexed listening effort (Peelle, 2018; Pichora-Fuller et al., 2016). In this literature, the magnitude of the task-evoked pupil response is typically measured during the stimulus and compared between conditions that differ in difficulty. For example, studies employ multiple levels of speech degradation or memory load (Alnæs et al., 2014; Zekveld et al., 2014). Therein, motivational factors are customarily neglected. This neglect has been widely acknowledged (Pichora-Fuller et al., 2016), but not addressed. In our sustained-attention value task, the perceptual difficulty (temporal coherence) was held constant across the entire session, whereas motivation (driven by task utility) was changed in blocks. Furthermore, we focus on the pre-stimulus, so-called 'tonic', pupil-linked arousal measured before each trial, rather than peri-stimulus, so-called 'phasic', stimulus-driven pupil dilation (de Gee et al., 2020). Future work should determine the interaction of these complementary arousal functions in behavior, perhaps by combining non-stationarities in both task utility and perceptual difficulty.

The specific pattern we observed in the dependence of performance on arousal, and its relation to task utility, likely illustrates both general principles as well as task- and species-specific patterns. For example, in contrast to our findings, a previous human study reported higher pre-trial pupil size during high reward blocks (Massar et al., 2016). This task was perceptually easy, while our sustained-attention value task was perceptually hard. An extensive literature shows that the relationship between tonic arousal and behavioral performance depends on task difficulty, with higher arousal being optimal for easier tasks (Sörensen et al., 2021; Yerkes & Dodson, 1908). The discrepancy between the findings reported by Massar et al. (2016) and ours might also be due to species difference; perhaps humans in laboratory conditions are on average in a lower arousal state than mice and thus typically sit on the opposite side of optimality. However, this is not the case for all individuals; moving closer to the optimal arousal state after increases in task utility involves either increases or decreases in arousal, depending on one's starting point (de Gee et al., 2020).

Taken together, our results add to the growing evidence for an inverted-U, three-state model for the role of arousal in behavioral performance (McGinley, David, et al., 2015; McGinley, Vinck, et al., 2015; Schriver et al., 2018; Yerkes & Dodson, 1908). This model accounts for the prominent exploration-exploitation tradeoff (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011) as the right side of the inverted-U. Pupil size was larger and mice walked more in low-reward blocks, a form of combined aroused and active exploration. In low-reward blocks, mice were also less likely to attend to and accumulate the relevant sensory evidence (temporal coherence). This result is in line with a recent observation that lapses in perceptual decisions reflect exploration (Pisupati et al., 2021). When pupil was smallest (below optimal levels) animal's exhibited rare and slow behavioral responses, indicative of a resting form of task disengagement. In sum, increases in task utility motivate mice to pay more attention, helped by the stabilization of pupil-linked arousal closer to its mid-sized, optimal level.

Methods

Animals

All surgical and animal handling procedures were carried out in accordance with the ethical guidelines of the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee (IACUC) of Baylor College of Medicine. A total of 114 animals were trained through to at least 5 sessions of the final phase of the task (see *Behavioral task*). We excluded 26 animals from the analysis (**Fig. S1P-R**) who had less than 5 sessions worth of data per animal after excluding sessions with an overall reward rate (see *Analysis and modeling of choice behavior*) of less than 15%. Thus, all remaining analyses are based on 88 mice (74 male, 14 female) aged 7-8 weeks at training onset. Wild-type mice were of C57BL/6 strain (Jackson Labs) (N=51; 1 female). Various heterozygous transgenic mouse lines used in this study were of Ai148 (IMSR Cat# JAX:030328; N=6; 3 females), Ai162 (IMSR Cat# JAX:031562; N=10, 3 females), ChAT-Cre (IMSR Cat# JAX:006410; N=3; all male), or ChAT-Cre crossed with Ai162 (N=18; 7 females). No differences were observed between genotypes or sexes, so results were pooled. This variety in genetic profile was required to target specific neural circuitries with two-photon imaging; the results of the imaging experiments are not reported, here. Mice received ad libitum water. Mice received ad libitum food on weekends but were otherwise placed on food restriction to maintain ~90% normal body weight. Animals were trained Monday-Friday. Mice were individually housed and kept on a regular light-dark cycle. All experiments were conducted during the light phase.

Head post implantation

The surgical station and instruments were sterilized prior to each surgical procedure. Isoflurane anesthetic gas (2–3% in oxygen) was used for the entire duration of all surgeries. The temperature of the mouse was maintained between 36.5°C and 37.5°C using a homoeothermic blanket system. After anesthetic induction, the mouse was placed in a stereotax (*Kopf Instruments*). The surgical site was shaved and cleaned with scrubs of betadine and alcohol. A 1.5-2 cm incision was made along the scalp mid-line, the scalp and overlying fascia were retracted from the skull. A sterile head post was then implanted using dental cement.

Behavioral task

Each 'trial' of the behavior consisted of three consecutive intervals (**Fig. 1A**): (i) the 'noise' or tone cloud interval, (ii) the 'signal' (temporal coherence) interval, and (iii) the inter-trial-interval (ITI). The duration of the noise interval was randomly drawn before each trial from an exponential distribution with mean of 5 seconds; this was done to ensure a flat hazard function for signal start time. In most sessions (82.8%), randomly drawn noise durations greater than 11 s were set to 11 s. In the remainder of sessions (17.2%), these trials were converted to a form of catch trial, consisting of 14 seconds of noise. Results were not affected by whether sessions included catch trials or not (**Fig. S2T-W**), and thus results were pooled for all further analyses. The duration of the signal interval was 3 s. The duration of the ITI was uniformly distributed between 2 and 3 s.

The noise stimulus was a 'tone cloud', consisting of consecutive chords of 20 ms duration (gated at start and end with 0.5 ms raised cosine). Each chord consisted of 12 pure tones,

selected randomly from a list of semitone steps from 1.5-96 kHz. For the signal stimulus, after the semi-random (randomly jittered by 1-2 semitones from tritone-spaced set of tones) first chord, all tones moved coherently upward by one semitone per chord. The ITI-stimulus was pink noise, which is highly perceptually distinct from the tone cloud. Stimuli were presented free field in the front left, upper hemifield at an overall intensity of 55 dB SPL (root-mean square [RMS]) using an intermittently calibrated Tucker Davis ES-1 electrostatic speakers and custom software system in LabVIEW.

Mice were head-fixed on a wheel and learned to lick for sugar water reward to report detection of the signal stimulus. Correct-go responses (hits) were followed by either 2 or 12 mL of sugar water, depending on block number. Reward magnitude alternated between 2 and 12 mL in blocks of 60 trials. Incorrect-go responses (false alarms) terminated the trial and were followed by a 14 s timeout with the same ITI-stimulus. Correct no-go responses (correct rejecting the full 14 s of noise) in the sessions that contained catch trails were also followed by 2 or 12 mL of sugar water in some sessions (see above).

Training mice to perform the sustained-attention value task involved three separate phases. In phase 1, the signal was louder than the noise sounds (58 and 52 dB, respectively), 'classical conditioning' trials (5 automatic rewards during the signal sounds) were included, and there were no block-based changes in reward magnitude (5 mL after every hit throughout the session). Phase 1 was conducted for four experimental sessions (**Fig. S1A**). In phase 2, we introduced the block-based changes in reward magnitude. Once mice obtained a reward rate higher than 25%, and the fraction of false alarm trials was below 50% for two out of three sessions in a row, they were moved up to the phase 3. Phase 2 lasted for 2 - 85 (median, 9) experimental sessions (**Fig. S1B**). Phase 3 was the final version of the task, with signal and noise stimuli of equal loudness and without any classical conditioning trials.

In a subset of experiments, the signal quality was systematically degraded by reducing the fraction of tones that moved coherently through time-frequency space. This is similar to reducing motion coherence in the classic random-dot motion task (Newsome et al., 1989). In these experiments, signal coherence was randomly drawn beforehand from six different levels: easy (100% coherence; as in the main task), hard (55-85% coherence), and four levels linearly spaced in between. In the main behavior, coherence ranged from 90-100% on each trial. Performance in the main behavior did not change with coherence level, so results were pooled across coherence levels.

After exclusion criteria (**Fig. S1P-R**), a total of 88 mice performed between 5 and 60 sessions (2100–24,960 trials per subject) of the final version of the sustained-attention value task (phase 3), yielding a total of 1983 sessions and 823,019 trials. A total of 10 mice performed the experiment with psychometrically degraded signals; they performed between 5 and 28 sessions (2083–11,607 trials per subject), yielding a total of 142 sessions and 58,826 trials.

Data acquisition

Custom LabVIEW software was written to execute the experiments, and synchronized all sounds, licks, pupil videography, and wheel motion. Licks were detected using a custom-made infrared beam-break sensor.

Pupil size. We continuously recorded images of the right eye with a Basler GigE camera (acA780-75gm), coupled with a fixed focal length lens (55 mm EFL, f/2.8, for 2/3"; Computar) and infrared filter (780 nm long pass; Midopt, BN810-43), positioned approximately 8 inches from the mouse. An off-axis infrared light source (two infrared LEDs; 850 nm, Digikey; adjustable in intensity and position) was used to yield high-quality images of the surface of the eye and a dark pupil. Images (504 × 500 pixels) were collected at 15 Hz, using a National Instruments PCIe-8233 GigE vision frame grabber and custom LabVIEW code. To achieve a wide dynamic range of pupil fluctuations, an additional near-ultraviolet LED (405-410 nm) was positioned above the animal and provided low intensity illumination that was adjusted such that the animal's pupil was approximately mid-range in diameter following placement of the animal in the set-up and did not saturate the eye when the animal walked.

Walking speed. We continuously measured treadmill motion using a rotary optical encoder (Accu, SL# 2204490) with a resolution of 8,000 counts/revolution.

Analysis and modeling of choice behavior

All analyses were performed using custom-made Python scripts, unless stated otherwise.

Trial exclusion criteria. We excluded the first (low reward) block of each session, as mice spent this block (termed block '0'; low reward) becoming engaged in the task (**Fig. S2A-C;** see also **Fig. S3B**). We found that a small fraction of trials began during a lick bout that had already started during the preceding ITI. These trials were immediately terminated. These rare 'false start trials' (2.5±0.2 % s.e.m. of trials across mice), were removed from further analyses. When pooling data across trials within a block, we always excluded the first 14 trials after the first hit in each block (in both high and low reward blocks; see also *Time course of behavioral adjustments*, below).

Behavioral metrics. Due to the quasi-continuous nature of the task, we could not use classic signal detection theory (Green & Swets, 1966) to compute sensitivity (d') and choice bias (criterion). As analogous measures (**Fig. S2D**), we computed 'discriminatory response rate' and 'overall response rate'. We defined discriminatory response rate as:

$$R_s = \frac{N_{HIT}}{T_S} - \frac{N_{FA}}{T_N} \qquad \text{Eq. 1}$$

where N_{HIT} is the number of hit trials, N_{FA} is the number of false alarm trials, T_S is the total duration a signal sound was played (in seconds) before licks and T_N the total time a noise sound was played. We defined overall response rate as:

$$R_o = \frac{N_{HIT} + N_{FA}}{T_S + T_N} \qquad \text{Eq. 2}$$

Reward rate was defined as the percentage of trials that ended in a hit (response during the signal). Reaction time on hit trials was defined as the time from signal onset until the response (first lick).

To characterize the theoretical relationship between discriminatory response rate and overall response rate (**Fig. S2D**), and to calculate the optimal overall response rate in our task (**Fig. S2E**), we generated a simulated data set. Specifically, we generated synthetic trials that matched the empirical trials (noise duration was drawn from an exponential distribution with mean = 5 s; truncated at 11 s). We then systematically varied overall response rate by drawing random response times from exponential distributions with various means (1 / rate) and assigned those (random) response times to the synthetic trials. We varied the overall response rate (rate) from 0 to 1 responses/s, in 100 evenly spaced steps. For each response rate, the decision agent performed 500K simulated trials. For each iteration we then calculated the resulting reward rate and discriminatory response rate.

Time course of learning. To characterize animal's learning, we fitted the following function:

$$B(s) = a \cdot e^{-b \cdot s} + c \qquad \text{Eq. 3}$$

where B is a behavioral metric of interest, s is session number with respect to start of phase 3, and a, b and c the free parameters of the fit.

Time course of behavioral adjustments. To calculate the speed of within-block behavioral adjustments to changes in task utility, we fitted the following function:

$$B(t) = a \cdot \ln(t) + b \cdot t + c \qquad \text{Eq. 4}$$

where B is a behavioral metric of interest, t is trial number since first correct response (hit) in a block, and a, b and c the free parameters of the fit. We then calculated the difference between the maximum and minimum of the fitted function and calculated the trial number for which 95% of this difference was reached. For overall response rate, discriminatory response rate and reward rate, this occurred on average at 15 trials after the first correct response in a low reward block (**Fig. 2B,E,H**). Therefore, when pooling data across trials within a block, we always excluded the first 14 trials after the first hit in each block (in both high and low reward blocks). We verified that our conclusions are not affected by specifics of this trial-selection procedure (**Fig. S2L-O**).

Accumulation-to-bound modeling. We fitted the choice and reaction time data with an accumulation-to-bound model of decision-making. The model was fitted to all data of all animals combined, but separately for each block number (**Fig. 5A-C**). In a separate analysis, we fitted the model separately for four different arousal states and for the high and low task utility blocks; we then averaged the fits across task utility to pinpoint the effect of arousal on decision-making that is independent of task-utility (**Fig. 5D-F**). The model was fitted based on continuous maximum

likelihood using the Python-package PyDDM (Shinn et al., 2020). The combination of all model parameters determines the fraction of correct responses and their associated RTs (**Fig. 4B-D**). However, their effects on the decision variable are distinct and can therefore be dissociated by simultaneously fitting choice fractions and associated reaction time distributions. We formulated a single accumulator model that describes the accumulation of noisy sensory evidence toward a single choice boundary for a go-response.

In the 'basic model' (**Fig. S4A,B**), the decision dynamics were governed by leak and gaussian noise during noise (tone cloud) stimuli, and additionally by the drift rate during the signal (temporal coherence) stimuli (**Fig. 4A**):

$$\Delta y = (-y \cdot k + s \cdot v) \cdot \Delta t + cdW \qquad \text{Eq. 5}$$

where y is the decision variable (black example trace in **Fig. 4A**, right), k is the leak and controls the effective accumulation time constant ($1/k$), s is the stimulus category (0 during noise sounds; 1 during signal sounds), v is the drift rate and controls the overall efficiency of accumulation of relevant evidence (coherence), and cdW is Gaussian distributed white noise with mean 0 and variance $c^2\Delta t$. Evidence accumulation terminated at bound height 1 (go response) or at the end of the trial (no-go response), whichever came first. The starting point of evidence accumulation was fixed to 0.

In the 'attention lapse model' (**Fig. S4C,D**), the decision dynamics were additionally governed by an attention lapse probability:

$$\Delta y = (-y \cdot k + s \cdot l \cdot v) \cdot \Delta t + cdW \qquad \text{Eq. 6}$$

where l is a Bernoulli trial, that determined with probability p the fraction of signals on which the relevant sensory evidence s was accumulated.

The 'attention lapse + starting point variability model' (**Fig. S4E,F**) was the same as the 'attention lapse model', but the starting point of evidence accumulation was additionally uniformly distributed between 0 and sz. Fast errors are typically accounted for by allowing starting point of evidence accumulation to be variable (Laming, 1968).

The 'full model' (**Fig. 4**) was the same as the 'attention lapse model', but included a mixture rate, a binomial probability that determined the fraction of trials on which the decision dynamics were not governed by equation 6, but instead on which reaction times were randomly drawn from a gaussian distribution with mean μ and variance σ (μ and σ were fixed across block number). A similar 'dual process' model was recently proposed for rats making perceptual decisions based on auditory information (Hernández-Navarro et al., 2021).

We additionally considered a version of the full model in which bound height could vary with block number, and leak was fixed at 0 (**Fig. S4G,H**).

To calculate the optimal leak in the sustained-attention value task (**Fig. S4J,K**), we generated a simulated data set. In our simulations, we generated synthetic trials that matched the empirical trials (noise duration was drawn from an exponential distribution with mean = 5 s; truncated at 11 s). We then systematically varied leak from 0 to 10, in steps of 0.2; the other

parameters were the same as estimated from the empirical data (**Fig. 5B,C** & **Fig. S5A,B**). In each iteration, we simulated a decision agent performing 500K trials. For each iteration, we then calculated the resulting discriminatory response rate and reward rate.

Analysis of pupil data

All analyses were performed using custom-made Python scripts, unless stated otherwise.

Preprocessing. We measured pupil size and exposed eye area from the videos of the animal's eye using DeepLabCut (Mathis et al., 2018; Mridha et al., 2021). In approximately 1000 training frames randomly sampled across all sessions, we manually identified 8 points spaced at approximately evenly around the pupil, and 8 points evenly spaced around the eyelids. The network (resnet 110) was trained with default parameters. To increase the network's speed and accuracy when labeling (unseen) frames of all videos, we specified video-wise cropping values in the DeepLabCut configuration file that corresponded to a square around the eye. The pupil size (or exposed eye area) was computed as the area of an ellipse fitted to the detected pupil (exposed eye) points. If two or more points were labeled with a likelihood smaller than 0.1 (e.g., during blinks), we did not fit an ellipse, but flagged the frame as missing data. We then applied the following signal processing to the pupil (exposed eye) time series of each measurement session: (i) resampling to 10 Hz; (ii) blinks were detected by a custom algorithm that marked outliers in the z-scored temporal derivative of the pupil size time series; (iii) linear interpolation of missing or poor data due to blinks (interpolation time window, from 150 ms before until 150 ms after missing data); (iv) low-pass filtering (third-order Butterworth, cut-off: 3 Hz); and (v) conversion to percentage of the 99.9 percentile of the time series. See (McGinley, David, et al., 2015; Mridha et al., 2021) for details.

Quantification of pre-trial pupil size. We quantified pre-trial pupil size as the mean pupil size during the 0.25 s before trial onset. Pre-trial pupil size was highest after previous hits (**Fig. S3A**), likely because the phasic lick-related pupil response did not have enough time to return to baseline. We thus removed (via linear regression) components explained by the choice (go vs. no-go) and reward (reward vs. no reward) on the previous trial. We obtained qualitatively identical results without removing trial-to-trial variations of previous choices and rewards from pre-trial pupil size measures (**Fig. S3M-P**). To capture how close the animal's arousal state on each trial was to the optimal level, we computed the absolute difference between each pre-trial's pupil size and the optimal size. Here, optimal size (28% of max) was defined as the pre-trial baseline pupil size for which discriminatory response rate was maximal (green vertical line in **Fig. 3B**).

Analysis of walking data

The instantaneous walking speed data was resampled to 10 Hz. We quantified pre-trial walking speed as the mean walking velocity during the 2 s before trial onset. We defined walking probability as the fraction of trials for which the absolute walking speed exceeded 1.25 cm/s (**Fig. S3C**).

Statistical comparisons

We used a one-way repeated measures ANOVA to test for the effect of signal coherence (**Fig. S1N,O**). We used a 3 × 2 repeated measures ANOVA to test for the main effect of task utility and time-on-task (block number of a given reward magnitude), and their interaction (**Fig. 2C,F,I** & **Fig. 3E**).

We directly compared bootstrapped distributions of the model parameter estimates to test for the main effects of task utility and time-on-task (**Fig. 6**). We used Bayesian information criterion (BIC) for model selection and verified whether the added complexity of each model was justified to account for the data (**Fig. S4I**). A difference in BIC of 10 is generally taken as a threshold for considering one model a sufficiently better fit (Spiegelhalter et al., 2002).

We used sequential polynomial regression analysis (Draper & Smith, 1998), to quantify whether the relationships between pre-trial pupil size and behavioral measures were better described by a 1st order (linear) up to a 5th order model (**Fig. 3B,C**):

$$\mathbf{Y} = \beta_0 1 + \beta_1 \mathbf{X} + \beta_2 \mathbf{X}^2 \ldots \qquad \text{Eq. 7}$$

where **Y** was a vector of the dependent variable (e.g., bin-wise discriminatory response rate), **X** was a vector of the independent variable (e.g. bin-wise pre-trial pupil size), and β as polynomial coefficients. To assess the amount of variance that each predictor accounted for independently, we orthogonalized the regressors prior to model fitting using QR-decomposition. Starting with the zero-order (constant) model and based on F-statistics (Draper & Smith, 1998), we tested whether incrementally adding higher-order predictors improves the model significantly (explains significantly more variance). We tested 1st up to 5th order models.

We used mediation analysis (Baron & Kenny, 1986) to characterize the interaction between task utility, arousal, and attention (**Fig. 4**). We fitted the following linear regression models based on standard mediation path analysis:

$$\mathbf{C} = i_0 1 + c\mathbf{R} \qquad \text{Eq. 8}$$

$$\mathbf{P} = i_1 1 + a\mathbf{R} \qquad \text{Eq. 9}$$

$$\mathbf{C} = i_2 1 + c'\mathbf{R} + b\mathbf{P} \qquad \text{Eq. 10}$$

where **C** was a vector of the block-wise discriminatory response rate, **R** was a vector of the block-wise reward context (0 for low reward; 1 for high reward), **P** was a vector of block-wise pre-trial pupil measures, and c, c', a, b, i_0, i_1 and i_2 were the free parameters of the fit. The parameters were fit using freely available Python software (Vallat, 2018).

All tests were performed two-tailed. All error bars are 68% bootstrap confidence intervals of the mean, unless stated otherwise.

Data availability

Data will be made publicly available upon publication.

Code availability
Analysis scripts will be made publicly available upon publication.

References

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R. F., Dayan, P., & Costa, R. M. (2021). The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron*, *109*(1), 149-163.e7. https://doi.org/10.1016/j.neuron.2020.10.013

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1. https://doi.org/10.1167/14.4.1

Arnsten, A. F. T., & Goldman-Rakic, P. S. (1984). Selective prefrontal cortical projections to the region of the locus coeruleus and raphe nuclei in the rhesus monkey. *Brain Research*, *306*(1), 9–18. https://doi.org/10.1016/0006-8993(84)90351-2

Arnsten, A. F. T., & Li, B.-M. (2005). Neurobiology of Executive Functions: Catecholamine Influences on Prefrontal Cortical Functions. *Biological Psychiatry*, *57*(11), 1377–1384. https://doi.org/10.1016/j.biopsych.2004.08.019

Ashwood, Z. C., Roy, N. A., Stone, I. R., Urai, A. E., Churchland, A. K., Pouget, A., & Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, *25*(2), Article 2. https://doi.org/10.1038/s41593-021-01007-z

Aston-Jones, G., & Cohen, J. D. (2005). An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Bari, A., Xu, S., Pignatelli, M., Takeuchi, D., Feng, J., Li, Y., & Tonegawa, S. (2020). Differential attentional control mechanisms by two distinct noradrenergic coeruleo-frontal cortical pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(46), 29080–29089. https://doi.org/10.1073/pnas.2015635117

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research. Brain Research Reviews*, *42*(1), 33–84.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765.

Botvinick, M., & Braver, T. (2015). Motivation and Cognitive Control: From Behavior to Neural Mechanism. *Annual Review of Psychology*, *66*(1), 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology*, *40*, 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545

Breton-Provencher, V., & Sur, M. (2019). Active control of arousal by a locus coeruleus GABAergic circuit. *Nature Neuroscience*, *22*(2), 218–228. https://doi.org/10.1038/s41593-018-0305-z

Brody, C. D., & Hanks, T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion in Neurobiology*, *37*, 149–157.

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, *340*(6128), 95–98. https://doi.org/10.1126/science.1233912

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Cazettes, F., Reato, D., Morais, J. P., Renart, A., & Mainen, Z. F. (2021). Phasic Activation of Dorsal Raphe Serotonergic Neurons Increases Pupil Size. *Current Biology*, *31*(1), 192-197.e4. https://doi.org/10.1016/j.cub.2020.09.090

de Gee, J. W., Colizoli, O., Kloosterman, N. A., Knapen, T., Nieuwenhuis, S., & Donner, T. H. (2017). Dynamic modulation of decision biases by brainstem arousal systems. *ELife*, *6*, 309.

de Gee, J. W., Tsetsos, K., Schwabe, L., Urai, A. E., McCormick, D., McGinley, M. J., & Donner, T. H. (2020). Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains. *ELife*, *9*, e54014. https://doi.org/10.7554/eLife.54014

de Lecea, L., Carter, M. E., & Adamantidis, A. (2012). Shining Light on Wakefulness and Arousal. *Biological Psychiatry*, *71*(12), 1046–1052. https://doi.org/10.1016/j.biopsych.2012.01.032

Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis*. Wiley-Interscience.

Engelmann, J. B., & Pessoa, L. (2014). Motivation sharpens exogenous spatial attention. *Motivation Science*, *1*(S), 64–72. https://doi.org/10.1037/2333-8113.1.S.64

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—Focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*(4), 437–455. https://doi.org/10.1016/j.conb.2007.07.011

Ghosh, S., & Maunsell, J. H. R. (2021). Single trial neuronal activity dynamics of attentional intensity in monkey visual area V4. *Nature Communications*, *12*(1), 2003. https://doi.org/10.1038/s41467-021-22281-2

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 252–269.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574.

Goris, R. L. T., Movshon, J. A., & Simoncelli, E. P. (2014). Partitioning neuronal variability. *Nature Neuroscience*, *17*(6), Article 6. https://doi.org/10.1038/nn.3711

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics. 1966*. New York.

Grossman, C. D., Bari, B. A., & Cohen, J. Y. (2022). Serotonin neurons modulate learning rate through uncertainty. *Current Biology*, *32*(3), 586-599.e7. https://doi.org/10.1016/j.cub.2021.12.006

Grujic, N., Brus, J., Burdakov, D., & Polania, R. (2022). Rational inattention in mice. *Science Advances*, *8*(9), eabj8935. https://doi.org/10.1126/sciadv.abj8935

Grujic, N., Tesmer, A., Bracey, E. F., Peleg-Raibstein, D., & Burdakov, D. (2022). *Control and coding of pupil size by hypothalamic orexin neurons* (p. 2022.04.12.488026). bioRxiv. https://doi.org/10.1101/2022.04.12.488026

Gutierrez-Castellanos, N., Sarra, D., Godinho, B. S., & Mainen, Z. F. (2022). *Maturation of prefrontal input to dorsal raphe increases behavioral persistence in mice* (p. 2022.01.01.474690). bioRxiv. https://doi.org/10.1101/2022.01.01.474690

Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J., & Berke, J. D. (2016). Mesolimbic dopamine signals the value of work. *Nature Neuroscience*, *19*(1), 117–126. https://doi.org/10.1038/nn.4173

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., & Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*.

Harris, K. D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience*, *12*(9), 509–523.

Hasselmo, M. E., & McGaughy, J. (2004). High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. In *Progress in Brain Research* (Vol. 145, pp. 207–231). Elsevier. https://doi.org/10.1016/S0079-6123(03)45015-2

Hernández-Navarro, L., Hermoso-Mendizabal, A., Duque, D., de la Rocha, J., & Hyafil, A. (2021). Proactive and reactive accumulation-to-bound processes compete during perceptual decisions. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-27302-8

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science (New York, N.Y.)*, *143*(3611), 1190–1192.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–1596.

Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2020.03.005

Joshi, S., & Gold, J. I. (2022). Context-dependent relationships between locus coeruleus firing patterns and coordinated neural activity in the anterior cingulate cortex. *ELife*, *11*, e63490. https://doi.org/10.7554/eLife.63490

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science (New York, N.Y.)*, *154*(3756), 1583–1585.

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual Deficit during a Mental Task. *Science (New York, N.Y.)*, *157*(3785), 218–219.

Kane, G. A., Vazey, E. M., Wilson, R. C., Shenhav, A., Daw, N. D., Aston-Jones, G., & Cohen, J. D. (2017). Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(6), 1073–1083. https://doi.org/10.3758/s13415-017-0531-y

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Academic Press.

Lee, S.-H., & Dan, Y. (2012). Neuromodulation of brain states. *Neuron*, *76*(1), 209–222.

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: Behavior, brain activation, and individual differences. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 99–112. https://doi.org/10.3758/CABN.8.1.99

Massar, S. A. A., Lim, J., Sasmita, K., & Chee, M. W. L. (2016). Rewards boost sustained attention through higher effort: A value-based decision making approach. *Biological Psychology*, *120*, 21–27. https://doi.org/10.1016/j.biopsycho.2016.07.019

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

McGinley, M. J., David, S. V., & McCormick, D. A. (2015). Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron*, *87*(1), 179–192.

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., Tolias, A. S., Cardin, J. A., & McCormick, D. A. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron*, *87*(6), 1143–1161.

Mridha, Z., de Gee, J. W., Shi, Y., Alkashgari, R., Williams, J., Suminski, A., Ward, M. P., Zhang, W., & McGinley, M. J. (2021). Graded recruitment of pupil-linked neuromodulation by parametric stimulation of the vagus nerve. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-21730-2

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154.

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), Article 10. https://doi.org/10.1038/s41593-019-0502-4

Nestvogel, D. B., & McCormick, D. A. (2022). Visual thalamocortical mechanisms of waking state-dependent activity and alpha oscillations. *Neuron*, *110*(1), 120-138.e4. https://doi.org/10.1016/j.neuron.2021.10.005

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, *341*(6237), 52–54. https://doi.org/10.1038/341052a0

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013). The Timescale of Perceptual Evidence Integration Can Be Adapted to the Environment. *Current Biology*, *23*(11), 981–986.

Peelle, J. E. (2018). Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. *Ear and Hearing*, *39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Pfeffer, T., Ponce-Alvarez, A., Tsetsos, K., Meindertsma, T., Gahnström, C. J., Brink, R. L. van den, Nolte, G., Engel, A. K., Deco, G., & Donner, T. H. (2021). Circuit mechanisms for the chemical modulation of cortex-wide network interactions and behavioral variability. *Science Advances*, *7*(29), eabf5620. https://doi.org/10.1126/sciadv.abf5620

Pfeiffer, B. E. (2020). The content of hippocampal "replay." *Hippocampus*, *30*(1), 6–18. https://doi.org/10.1002/hipo.22824

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, *37*, 5S. https://doi.org/10.1097/AUD.0000000000000312

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., & Churchland, A. K. (2021). Lapses in perceptual decisions reflect exploration. *ELife*, *10*, e55490. https://doi.org/10.7554/eLife.55490

Polack, P.-O., Friedman, J., & Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nature Neuroscience*, *16*(9), 1331–1339.

Porrino, L. J., & Goldman-Rakic, P. S. (1982). Brainstem innervation of prefrontal and anterior cingulate cortex in the rhesus monkey revealed by retrograde transport of HRP. *Journal of Comparative Neurology*, *205*(1), 63–76. https://doi.org/10.1002/cne.902050107

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, *7*, 13289.

Reinagel, P. (2021). Rational regulation of water-seeking effort in rodents. *Proceedings of the National Academy of Sciences*, *118*(48). https://doi.org/10.1073/pnas.2111742118

Richter, M., Gendolla, G. H. E., & Wright, R. A. (2016). Chapter Five - Three Decades of Research on Motivational Intensity Theory: What We Have Learned About Effort and What We Still Don't Know. In A. J. Elliot (Ed.), *Advances in Motivation Science* (Vol. 3, pp. 149–186). Elsevier. https://doi.org/10.1016/bs.adms.2016.02.001

Robbins, T. (2002). The 5-choice serial reaction time task: Behavioural pharmacology and functional neurochemistry. *Psychopharmacology*, *163*(3), 362–380. https://doi.org/10.1007/s00213-002-1154-7

Robert, B., Kimchi, E. Y., Watanabe, Y., Chakoma, T., Jing, M., Li, Y., & Polley, D. B. (2021). A functional topography within the cholinergic basal forebrain for encoding sensory cues and behavioral reinforcement outcomes. *ELife*, *10*, e69514. https://doi.org/10.7554/eLife.69514

Sander, C., Hensch, T., Wittekind, D. A., Böttger, D., & Hegerl, U. (2015). Assessment of Wakefulness and Brain Arousal Regulation in Psychiatric Research. *Neuropsychobiology*, *72*(3–4), 195–205. https://doi.org/10.1159/000439384

Sarter, M., Gehring, W. J., & Kozak, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews*, *51*(2), 145–160. https://doi.org/10.1016/j.brainresrev.2005.11.002

Schriver, B. J., Bagdasarov, S., & Wang, Q. (2018). Pupil-linked arousal modulates behavior in rats performing a whisker deflection direction discrimination task. *Journal of Neurophysiology*, *120*(4), 1655–1670. https://doi.org/10.1152/jn.00290.2018

Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Neuron*, *80*(3), 791–806.

Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neurosciences*, *34*(3), 114–123. https://doi.org/10.1016/j.tins.2010.11.002

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *BioRxiv*, 2020.03.14.992065. https://doi.org/10.1101/2020.03.14.992065

Siegel, M., Engel, A. K., & Donner, T. H. (2011). Cortical network dynamics of perceptual decision-making in the human brain. *Frontiers in Human Neuroscience*, *5*, 21.

Sörensen, L. K. A., Bohté, S. M., Slagter, H. A., & Scholte, H. S. (2021). *Deep hierarchical sensory processing accounts for effects of arousal state on perceptual decision-making* (p. 2021.05.19.444798). https://doi.org/10.1101/2021.05.19.444798

Spiegelhalter, D. J., Best, N. G., & Carlin, B. P. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal \ldots*.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory Consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, *576*(7786), Article 7786. https://doi.org/10.1038/s41586-019-1787-x

Totah, N. K., Neves, R. M., Panzeri, S., Logothetis, N. K., & Eschenko, O. (2018). The Locus Coeruleus Is a Complex and Differentiated Neuromodulatory System. *Neuron*, *99*(5), 1055-1068.e6. https://doi.org/10.1016/j.neuron.2018.07.037

Tremblay, L., & Schultz, W. (2000). Reward-Related Neuronal Activity During Go-Nogo Task Performance in Primate Orbitofrontal Cortex. *Journal of Neurophysiology*, *83*(4), 1864–1876. https://doi.org/10.1152/jn.2000.83.4.1864

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The Role of Locus Coeruleus in the Regulation of Cognitive Performance. *Science*. https://doi.org/10.1126/science.283.5401.549

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592.

Vallat, R. (2018). Pingouin: Statistics in Python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, *360*(6388), 537–542. https://doi.org/10.1126/science.aar7186

Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and Dopamine Neurons in the Reward/Effort Trade-Off: A Direct Electrophysiological Comparison in Behaving Monkeys. *Journal of Neuroscience*, *35*(20), 7866–7877.

Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*, *16*(5), 639–647. https://doi.org/10.1038/nn.3377

Wang, L., & Krauzlis, R. J. (2018). Visual Selective Attention in Mice. *Current Biology*, *28*(5), 676-685.e4. https://doi.org/10.1016/j.cub.2018.01.038

Wang, X.-J. (2008). Decision making in recurrent neuronal circuits. *Neuron*, *60*(2), 215–234.

Weilnhammer, V., Stuke, H., Eckert, A.-L., Standvoss, K., & Sterzer, P. (2021). *Humans and mice fluctuate between external and internal modes of sensory processing* (p. 2021.08.20.457079). https://doi.org/10.1101/2021.08.20.457079

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482.

Yüzgeç, Ö., Prsa, M., Zimmermann, R., & Huber, D. (2018). Pupil Size Coupling to Cortical States Protects the Stability of Deep Sleep via Parasympathetic Modulation. *Current Biology*, *28*(3), 392-400.e3. https://doi.org/10.1016/j.cub.2017.12.049

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*, *101*, 76–86. https://doi.org/10.1016/j.neuroimage.2014.06.069

Zhao, S., Liu, Y., & Wei, K. (2022). Pupil-Linked Arousal Response Reveals Aberrant Attention Regulation among Children with Autism Spectrum Disorder. *Journal of Neuroscience*, *42*(27), 5427–5437. https://doi.org/10.1523/JNEUROSCI.0223-22.2022

Acknowledgements

We thank Daeyeol Lee for help developing the behavioral paradigm, Anton Banta for technical assistance, Sarim Aleem for help with pupil size analysis, many Rice University undergraduate students for help with animal behavioral training, Max Shinn, Konstantinos Tsetsos and Peter Murphy for helpful discussions about the drift diffusion modeling and Nelson Totah for helpful comments on a draft of the manuscript.

Author Contributions

Conceptualization: JWG, MJM
Funding Acquisition: MJM
Data acquisition software: WZ, HJ
Investigation: ZM, MH, YS, HR, SS, NK, MT, KJ
Data Curation: JWG, MH, YS, HR, SS, NK, MT, KJ
Formal analysis: JWG, MJM
Methodology: JWG, MJM
Writing—original draft: JWG, MJM
Writing—review and editing: JWG, ZM, MJM
Supervision: MJM
Project Administration: MJM

Competing Interests

The authors declare no competing interests.

Supplementary figures S1-S5

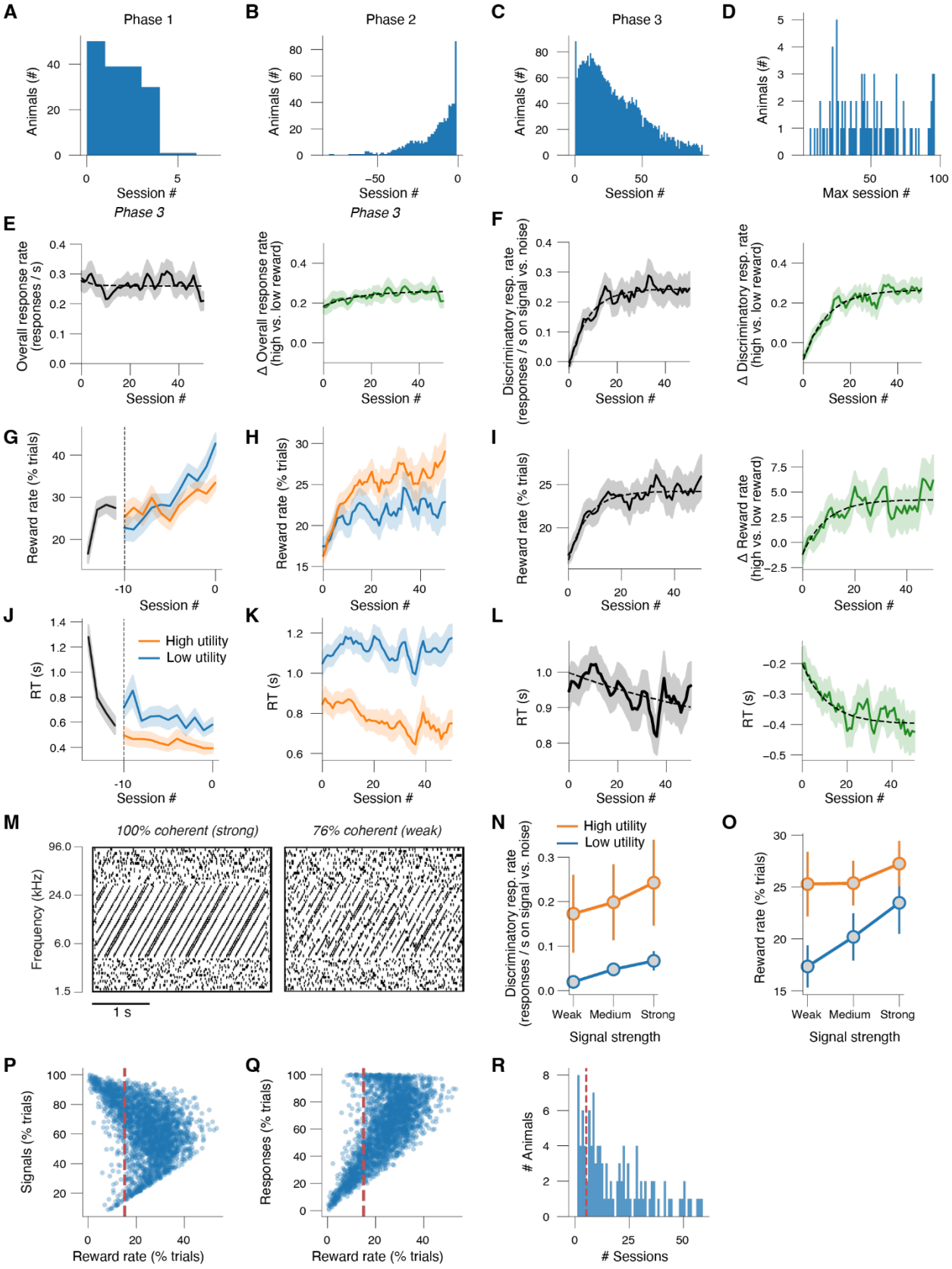
A
Phase 1
Animals (#)
Session #
B
Phase 2
Animals (#)
Session #
C
Phase 3
Animals (#)
Session #
D
Animals (#)
Max session #
E
Phase 3
Overall response rate (responses / s)
Session #
Phase 3
Δ Overall response rate (high vs. low reward)
Session #
F
Discriminatory resp. rate (responses / s on signal vs. noise)
Session #
Δ Discriminatory resp. rate (high vs. low reward)
Session #
G
Reward rate (% trials)
Session #
H
Reward rate (% trials)
Session #
I
Reward rate (% trials)
Session #
Δ Reward rate (high vs. low reward)
Session #
J
High utility
Low utility
RT (s)
Session #
K
RT (s)
Session #
L
RT (s)
Session #
RT (s)
Session #
M
100% coherent (strong)
76% coherent (weak)
Frequency (kHz)
1 s
N
High utility
Low utility
Discriminatory resp. rate (responses / s on signal vs. noise)
Weak
Medium
Strong
Signal strength
O
Reward rate (% trials)
Weak
Medium
Strong
Signal strength
P
Signals (% trials)
Reward rate (% trials)
Q
Responses (% trials)
Reward rate (% trials)
R
Animals
Sessions

Figure S1. We fitted exponential functions to quantify the speed of learning in phase 3 (Methods). The fitted time constants for discriminatory response rate and reward rate collapsed across reward blocks were $\tau = 8.0 \pm 0.7$ and $\tau = 7.2 \pm 1.1$, respectively (panels F,I, left). We verified that these time constants were not sensitive to sampling bias (dropout for later sessions) by refitting to strata of mice defined by the number sessions they performed (panel D). We also analyzed if and how rapidly mice learn to strategically regulate attentional effort. We found that, after extensive training, overall response rate, discriminatory response rate, and reward rate were all higher in the high compared to low reward blocks (Fig. 1D,F and panel Fig. S1H, orange vs. blue); reaction times (RTs) were lower (panel K). Thus, mice adaptively turned up their feature-based attention after increases in task utility. The fitted time constant for high vs. low reward discriminatory response rate and reward rate were $\tau = 10.0 \pm 1.5$ and $\tau = 9.2 \pm 4.2$, respectively (panels F,I, right), and not statistically different from the time constants for overall discriminatory response rate and reward rate ($p = 0.15$ and $p = 0.32$, respectively). Thus, mice learned to strategically allocate attention roughly at the same pace as overall perceptual learning. **(A)** Histogram of experimental session number in learning phase 1. **(B)** Histogram of experimental session number in learning phase 2 (with respect to last session number in phase 2). **(C)** Histogram of experimental session number in phase 3. **(D)** Histogram of the maximum session number per animal. **(E)** Overall response rate (Methods) across experimental sessions in learning phases 1 and 2; session numbers are with respect to the last session in phase 2. **(F)** Overall response rate across experimental sessions in learning phase 3, separately for high and low reward blocks. **(G)** Left: As F, but collapsed across reward context. Right: As F, but for the difference between high and low reward blocks. Dashed lines, exponential fit (Methods). **(H-J)** As E-G, but for discriminatory response rate (Methods). **(L-M)** As E-G, but for reward rate (Methods). **(N-P)** As E-G, but for reaction time on hit trials (Methods). **(Q)** Example spectrograms of strong (left) and weak (right) signals (Methods). **(R)** Discriminatory response rate (Methods) across three difficulty bins and separately for high and low reward blocks. Stats, 2-way repeated measures ANOVA (factors task utility [high vs. low] and signal strength (coherence) bin [weak, medium, strong]); main effect task utility: $F_{1,9} = 4.6$, $p = 0.060$; main effect signal strength: $F_{2,18} = 5.0$, $p = 0.018$; interaction effect: $F_{2,18} = 0.4$, $p = 0.669$. Error bars, 68% confidence interval across animals (N=10, n=142 sessions). **(S)** As R, but for reward rate (Methods). Main effect task utility: $F_{1,9} = 16.5$, $p = 0.003$; main effect signal strength: $F_{2,18} = 14.0$, $p < 0.001$; interaction effect: $F_{2,18} = 2.0$, $p = 0.162$. **(T)** Fraction of trials containing a signal plotted against reward rate (after excluding sessions in panel E). Every data point is a unique session. We excluded 455 sessions with a reward rate smaller than 15%, thereby excluding 4 animals. **(V)** As T, but for fraction of trial on which the animal responded on the y-axis. **(W)** Histogram of number of sessions per animal (after excluding sessions in panel A,B). We excluded 22 additional animals with fewer than 5 remaining sessions.

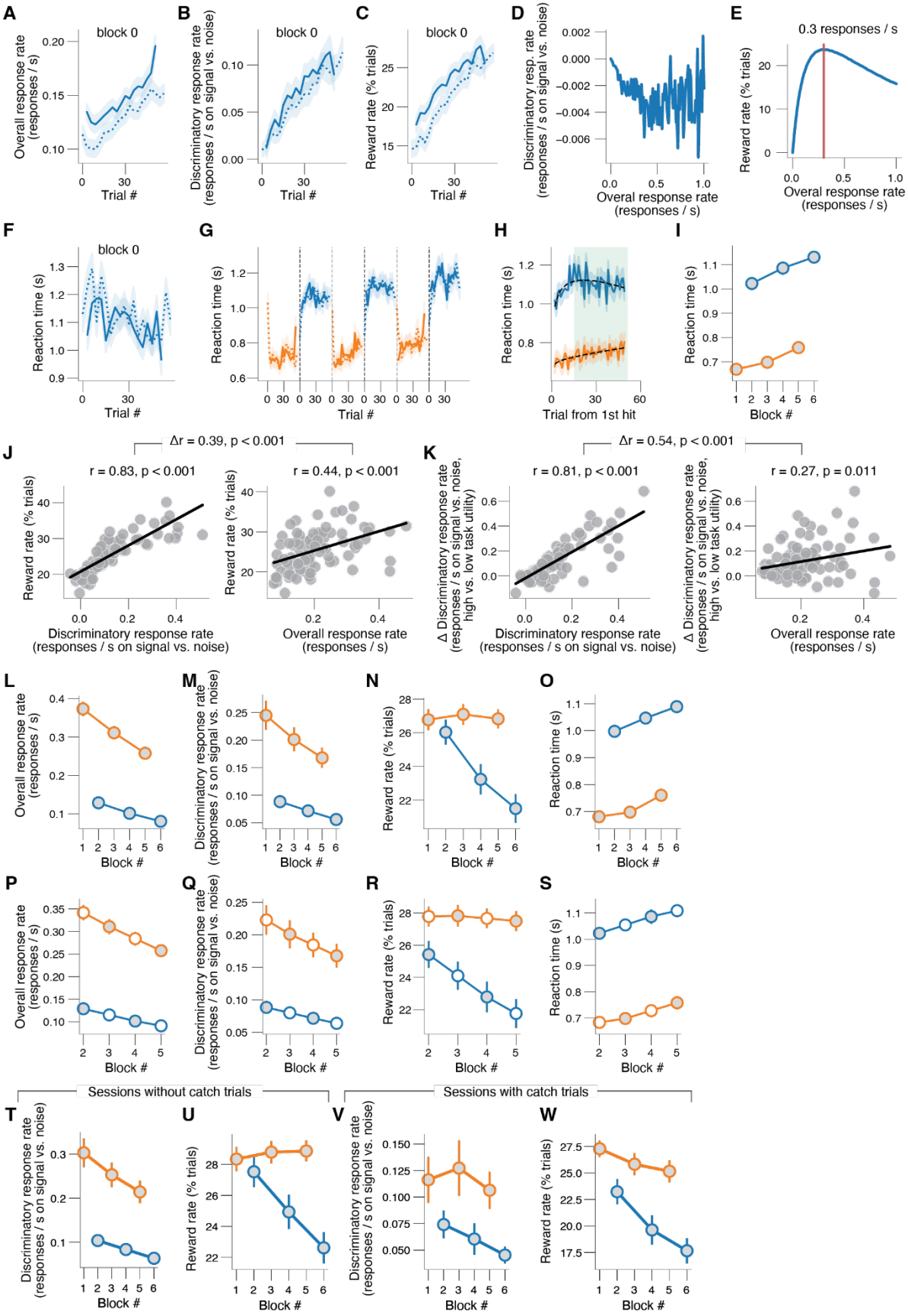
A
block 0
Overall response rate (responses / s)
Trial #
B
block 0
Discriminatory response rate (responses / s on signal vs. noise)
Trial #
C
block 0
Reward rate (% trials)
Trial #
D
Discriminatory resp. rate (responses / s on signal vs. noise)
Overal response rate (responses / s)
E
0.3 responses / s
Reward rate (% trials)
Overal response rate (responses / s)
F
block 0
Reaction time (s)
Trial #
G
Reaction time (s)
Trial #
H
Reaction time (s)
Trial from 1st hit
I
Reaction time (s)
Block #
Δr = 0.39, p < 0.001
J
r = 0.83, p < 0.001
r = 0.44, p < 0.001
Reward rate (% trials)
Discriminatory response rate (responses / s on signal vs. noise)
Overall response rate (responses / s)
Δr = 0.54, p < 0.001
K
r = 0.81, p < 0.001
r = 0.27, p = 0.011
Δ Discriminatory response rate (responses / s on signal vs. noise, high vs. low task utility)
Discriminatory response rate (responses / s on signal vs. noise)
Overall response rate (responses / s)
L
Overall response rate (responses / s)
Block #
M
Discriminatory response rate (responses / s on signal vs. noise)
Block #
N
Reward rate (% trials)
Block #
O
Reaction time (s)
Block #
P
Overall response rate (responses / s)
Block #
Q
Discriminatory response rate (responses / s on signal vs. noise)
Block #
R
Reward rate (% trials)
Block #
S
Reaction time (s)
Block #
Sessions without catch trials
Sessions with catch trials
T
Discriminatory response rate (responses / s on signal vs. noise)
Block #
U
Reward rate (% trials)
Block #
V
Discriminatory response rate (responses / s on signal vs. noise)
Block #
W
Reward rate (% trials)
Block #

Figure S2. (A) Overall response rate in the first (low reward) block of the session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(B)** As A, but for discriminatory response rate (Methods). **(C)** As A, but for reward rate (Methods). **(D)** Simulation of relationship between discriminatory response rate and overall response rate (Methods). **(E)** Simulation of relationship between reward rate and overall response rate (Methods). Red line, optimal overall response rate (reward rate peaks). **(F)** As A, but for reaction time (RT) on hit trials. **(G)** Reaction time (RT) on hit trials (Methods) across three high and three low reward blocks in the same experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(H)** As G, but collapsed across blocks of same reward magnitude. Green window, trials used when pooling data across trials within a block (as in panel I). **(I)** As G, but collapsed across trials within a block. Horizontal red line, optimal overall response rate (Methods). Stats, 2-way repeated measures ANOVA (factors task utility [high vs. low] and time-on-task [1, 2, 3]); main effect task utility: $F_{1,87} = 571.1$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 25.8$, $p < 0.001$; interaction effect: $F_{2,174} = 0.9$, $p = 0.401$. **(J)** Left: Mean reward rate (irrespective of task utility) plotted against mean discriminatory response rate (irrespective of task utility). Every data point is a unique session. Right: As left, but for mean overall response rate (irrespective of task utility) on the x-axis. We used multiple linear regression to test if mean discriminatory response rate and mean overall response rate significantly predicted mean reward rate. The overall regression was statistically significant ($R^2 = 0.690$, $F(2, 85) = 94.4$, $p < 0.001$). Mean discriminatory response rate but not mean overall response rate significantly predicted mean reward rate ($t = 11.65$, $p < 0.001$ and $t = -0.45$, $p = 0.655$, respectively). **(K)** Left: Change in discriminatory response rate (high vs low task utility) plotted against mean discriminatory response rate (irrespective of task utility). Every data point is a unique session. Right: As left, but for mean overall response rate (irrespective of task utility) on the x-axis. We used multiple linear regression to test if mean discriminatory response rate and mean overall response rate significantly predicted the change in discriminatory response rate after increases in task utility. The overall regression was statistically significant ($R^2 = 0.707$, $F(2, 85) = 102.6$, $p < 0.001$). Both mean discriminatory response rate and mean overall response rate significantly predicted the change in discriminatory response rate after increases in task utility ($t = 13.58$, $p < 0.001$ and $t = -3.77$, $p < 0.001$, respectively). **(L-O)** As panel I and main Fig. 2C,F,I, but when using all trials within a block (Methods). Main effects task utility are as follows: Panel K: $F_{1,87} = 317.3$, $p < 0.001$. Panel L: $F_{1, 87} = 63.0$, $p < 0.001$. Panel M: $F_{1, 87} = 14.5$, $p < 0.001$. Panel N: $F_{1, 87} = 638.3$, $p < 0.001$. **(P-S)** As panel I and main Fig. 2C,F,I, but when controlling for time on task. The white circles are the average of the two adjacent blocks of the same task utility. Main effects task utility are as follows: Panel O: $F_{1, 87} = 274.0$, $p < 0.001$. Panel P: $F_{1, 87} = 55.8$, $p < 0.001$. Panel Q: $F_{1, 87} = 17.9$, $p < 0.001$. Panel R: $F_{1, 87} = 411.6$, $p < 0.001$. **(T,U)** As main Fig. 2F,I, but only for 82.8% of sessions without catch trials (Methods). Main effects task utility are as follows: Panel S: $F_{1,69} = 58.9$, $p < 0.001$. Panel T: $F_{1, 69} = 9.6$, $p = 0.003$. **(V,W)** As main Fig. 2F,I, but only for 17.2% of sessions with catch trials (Methods). Main effects task utility are as follows: Panel U: $F_{1,43} = 15.0$, $p < 0.001$. Panel V: $F_{1, 43} = 22.9$, $p < 0.001$. All panels: shading or error bars, 68% confidence interval across animals (N=88; n=1983 sessions).

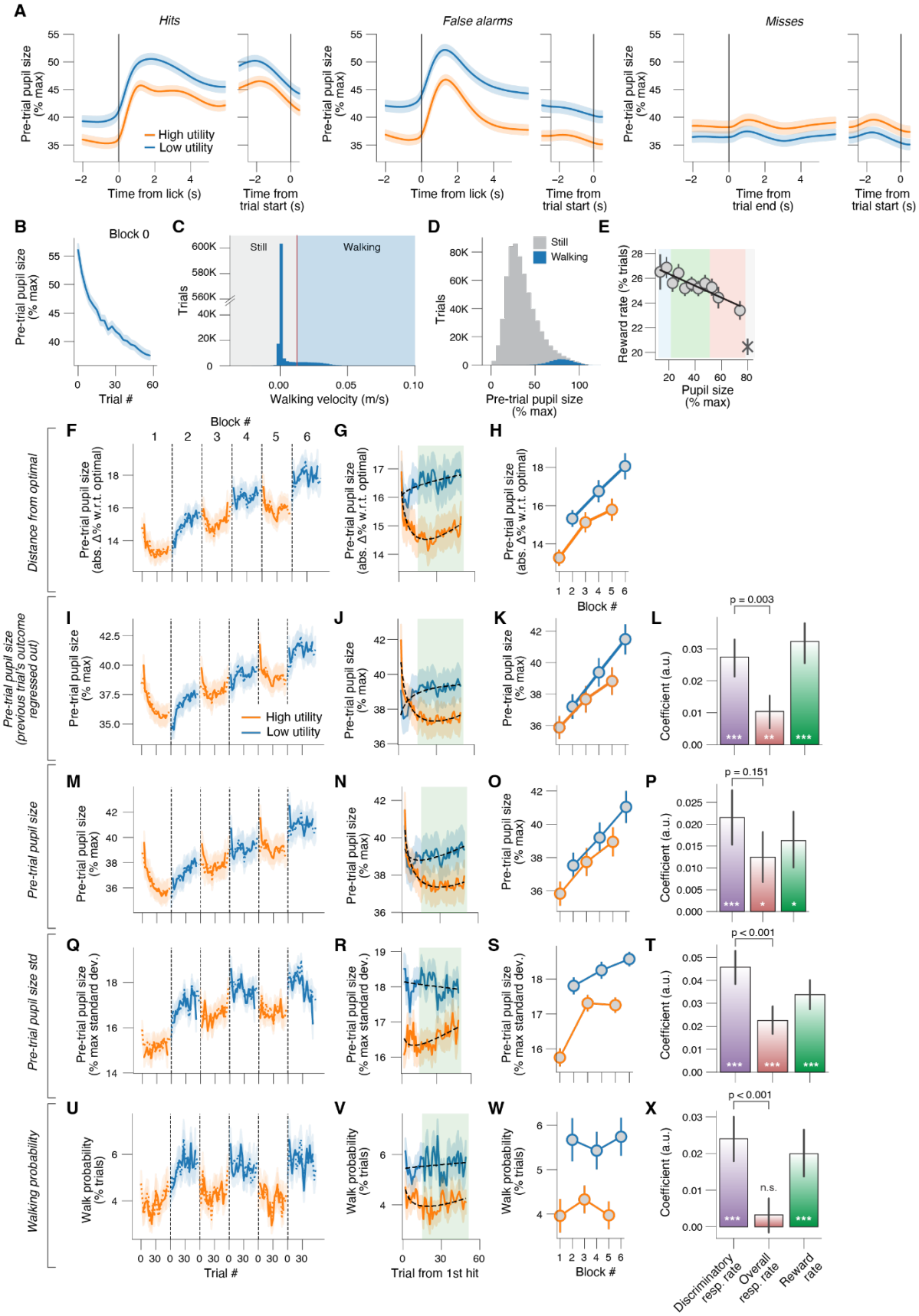
A
Hits
False alarms
Misses
Pre-trial pupil size (% max)
High utility
Low utility
Time from lick (s)
Time from trial start (s)
Time from trial end (s)
B
Block 0
Trial #
C
Still
Walking
Trials
Walking velocity (m/s)
D
Pre-trial pupil size (% max)
E
Reward rate (% trials)
Pupil size (% max)
Distance from optimal
Pre-trial pupil size (previous trial's outcome regressed out)
Pre-trial pupil size
Pre-trial pupil size std
Walking probability
Block #
Pre-trial pupil size (abs. Δ% w.r.t. optimal)
Pre-trial pupil size (% max standard dev.)
Walk probability (% trials)
Trial from 1st hit
Coefficient (a.u.)
p = 0.003
p = 0.151
p < 0.001
n.s.
Discriminatory resp. rate
Overall resp. rate
Reward rate

Figure S3. (A) Left: Pupil size across time on hit trials, separately for high and low reward blocks. Left, locked to licks; right, locked to next trial's onset. Middle, as left, but on false alarm trials. Right: As left, but for miss trials. Left, locked to trial offset (end of 3-s signal sound). **(B)** Pre-trial pupil size in the first (low reward) block of the session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(C)** Histogram of pre-trial walking velocity (across all animals and experimental sessions; Methods). Red line, cutoff for defining walking. **(D)** Histogram of pre-trial pupil size (across all animals and experimental sessions), separately for still and walking trials. **(E)** As main Fig. 3A, but for reward rate (Methods). 1st order fit: $F_{1,163} = 5.5$, $p = 0.020$; higher order models were not superior. **(F)** Absolute distance from optimal pre-trial pupil size across three high and three low reward blocks in the same experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(G)** As F, but collapsed across blocks of same reward magnitude. **(H)** As F, but collapsed across trials within a block. Stats, 2-way repeated measures ANOVA (factors task utility [high vs. low] and time-on-task [1, 2, 3]); main effect task utility: $F_{1,87} = 106.9$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 46.0$, $p < 0.001$; interaction effect: $F_{2,174} = 2.6$, $p = 0.075$. **(I)** Pre-trial pupil size (Methods) across three high and three low reward blocks in the same experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(J)** As J, but collapsed across blocks of same reward magnitude. **(K)** As I, but collapsed across trials within a block. Main effect task utility: $F_{1,87} = 41.8$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 51.0$, $p < 0.001$; interaction effect: $F_{2,174} = 6.3$, $p = 0.002$. **(L)** With pre-trial pupil size as mediation, fitted regression coefficients of the indirect path ($a \times b$; mediation). Stats, Wilcoxon signed-rank test; *, $p < 0.05$; ***, $p < 0.001$. **(M-P)** As I-L, but for pre-trial pupil size without previous trial's outcome regressed out (Methods). Stats in panel Q: 2-way repeated measures ANOVA (factors task utility [high vs. low] and time-on-task [1, 2, 3]); main effect task utility: $F_{1,87} = 35.2$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 42.9$, $p < 0.001$; interaction effect: $F_{2,174} = 1.3$, $p = 0.283$. **(Q-T)** As I-L, but for standard deviation of pre-trial pupil size. Stats in panel T: main effect task utility: $F_{1,87} = 108.8$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 25.3$, $p < 0.001$; interaction effect: $F_{2,174} = 8.4$, $p < 0.001$. **(U-X)** As I-L, but for walk probability (Methods). Stats in panel W: main effect task utility: $F_{1,87} = 53.2$, $p < 0.001$; main effect time-on-task: $F_{2,174} = 0.1$, $p = 0.935$; interaction effect: $F_{2,174} = 3.6$, $p = 0.029$. All panels: shading or error bars, 68% confidence interval across animals (N=88; n=1983 sessions).

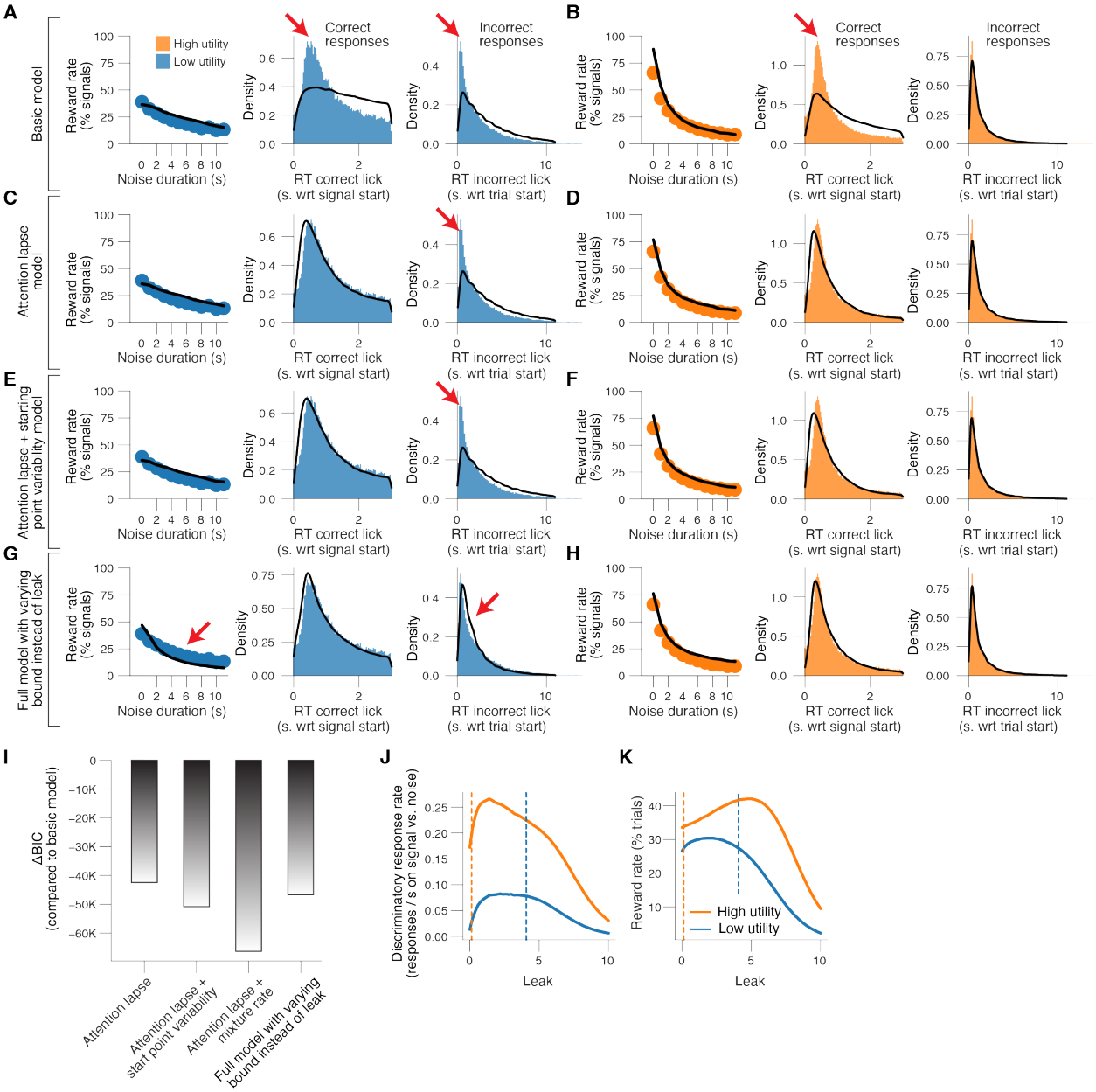

Figure S4. (A) Top: simulated overall response rate as a function of bound height (*a*) and leak (*k*) (Methods). Remaining parameters match those estimated based on the empirical data collapsed across the high and low reward blocks. Bottom: simulated overall response rate as a function attention lapse probability (*p*) and drift rate (*v*). Remaining parameters match those estimated based on the empirical data collapsed across the high and low reward blocks. **(B-D)** As A, but for discriminatory response rate, reward rate and RT on hit trials, respectively. **(E)** Left: Reward rate (defined as % correct responses on all trials containing signal sounds) in the low reward blocks, separately for binned noise durations. Middle: RT distribution for correct responses (hits) in the low reward blocks. Right: RT distribution for incorrect responses (false alarms) in the low reward blocks. Black lines, 'basic model' fit. In the 'basic model', the decision dynamics were governed by leak and gaussian noise during noise stimuli, and additionally by the drift rate during the signal stimuli (Methods). Red arrows, the 'basic model' failed to capture the shape of the RT

distribution associated with correct responses and failed to capture early RTs associated with incorrect responses in the low reward blocks. **(F)** As E, but for the high reward blocks. **(G,H)** As E,F, but for the 'attention lapse model'. The 'attention lapse model' was the same as the basic model but included an attention lapse probability (Methods). Red arrow, the 'attention lapse model' failed to capture early RTs associated with incorrect responses in the low reward blocks. **(I,J)** As E,F, but for 'attention lapse + starting point variability model'. The 'attention lapse + starting point variability model' was the same as the 'attention lapse model', but the starting point of evidence accumulation was not fixed at 0 but was drawn from a uniform distribution (Methods). Red arrow, the 'attention lapse + starting point variability model' still failed to fully capture early RTs associated with incorrect responses in the low reward blocks. **(K)** We compared the BIC between the four different models (Methods). The BIC for the 'basic model' was used as a baseline for each dataset. Lower BIC values indicate a model that is better able to explain the data, considering the model complexity; a difference BIC of 10 is generally taken as a threshold for considering one model a sufficiently better fit (Spiegelhalter et al., 2002). Formal model comparison based on Bayesian information criterion (BIC) favored the full model. **(L)** Simulated discriminatory response rate as a function of leak (k), separately for two sets of parameters that match those estimated based on the empirical data in the high and low reward blocks. Orange and blue dashed lines, estimated leak in the high and low reward blocks, respectively. **(M)** As L, but for simulated reward rate.

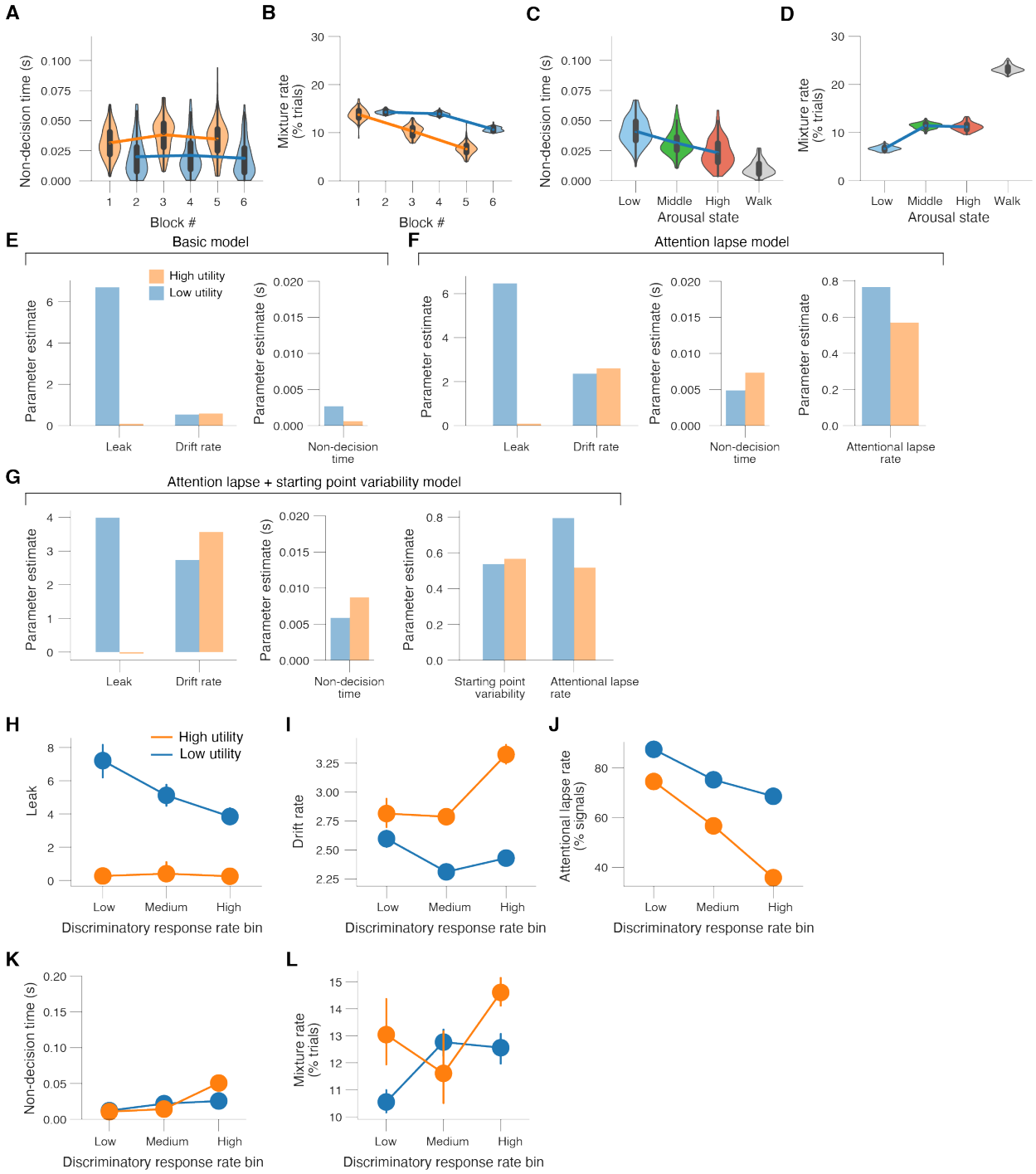

Figure S5. (A) Fitted non-decision time estimates (kernel density estimate of 100 bootstrapped replicates) separately per block number. Main effect task utility (fraction of bootstrapped parameter estimates in the low reward blocks higher than in the high reward blocks): p = 0.23. Main effect time-on-task (fraction of bootstrapped parameter estimates in the first two blocks higher than in the last two blocks): p = 0.44. **(B)** As A, but for mixture rate. Main effect task utility: p = 0.12. Main effect time-on-task: p = 0.03. **(C)** Fitted non-decision time estimates (kernel density estimate of 100 bootstrapped replicates) separately per arousal state (pupil size defined bins corresponding to colors in Fig. 3A-C; irrespective of task utility; Methods). No significant

differences between arousal states. **(D)** As C, but for mixture rate. The mixture rate was significantly lower during low vs middle arousal ($p < 0.01$) and during high arousal vs walking ($p < 0.01$). **(E)** Parameter estimates from the 'basic model' (Methods), separately per block type. **(F)** As E, but for the 'attention lapse' model (Methods). **(G)** As E, but for the 'attention lapse + starting point variability model' (Methods). **(H-L)** Parameter estimates from the 'full model' (mean and 68% confidence intervals across 25 bootstrapped replicates), now separately for the high and low reward blocks and separately for three discriminatory response rate defined bins. The three bins contain data of N=30, N=29 and N=29 mice with the lowest, middle and highest mean discriminatory response (irrespective of task utility), respectively.