Mice regulate their attentional intensity and arousal to exploit increases in task utility

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], 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

Because attention is limited in capacity and costly to utilize, organisms must continuously match their attentional intensity to its utility. Autonomic arousal strongly influences attention, but classical models give conflicting accounts of whether peak attentional intensity occurs during low, intermediate, or high arousal states. Furthermore, it is unclear if organisms can regulate their arousal to increase attention when it is useful. Here, we present results in mice from an auditory feature-based attention task with nonstationary task utility and use pupil size to infer arousal. We sample a wide range of arousal states by tracking pupil and behavior across 1.8 million trials from 88 mice. We show that peak attentional intensity occurs at intermediate arousal. Critically, mice stabilize their arousal near optimality when task utility is high. These results establish a strong and specific relationship between attention and arousal and demonstrate that self-regulation of arousal implements strategic attentional intensity allocation.

INTRODUCTION

Attention is limited in capacity and costly to utilize. Therefore, organisms are driven by ongoing behavioral incentives to choose *which* stimuli to attend to, and *how much* to attend to them. Daniel Kahneman termed these the selective and intensive aspects of attention, respectively (Kahneman, 1973). Characterizing the selective aspect of attention has been a cornerstone of systems neuroscience (Carrasco, 2011; Fritz et al., 2007; Maunsell & Treue, 2006). However, attentional intensity has received comparatively little scrutiny. Recent work emphasizes the importance of motivational factors in driving attentional intensity (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.

Manipulation of task utility in perceptual decision-making tasks provides a powerful approach to study the adaptive recruitment of attentional intensity. Since expectation of reward is a strong motivator (Roesch & Olson, 2004; Schultz, 2000), changes in task utility should lead to

corresponding changes in attentional intensity. In conventional task designs, signal detection-theoretic sensitivity provides a readout of attentional intensity. In line with these ideas, heightened reward expectation increases perceptual sensitivity (Engelmann & Pessoa, 2014; Ghosh & Maunsell, 2021) and also reduces reaction times (Locke & Braver, 2008) in humans and non-human primates. Rodents can match their rate of learning to the statistics of a dynamic environment (Grossman et al., 2022), perform cost-benefit analysis (Reinagel, 2021) and adapt their response vigor to task utility (Wang et al., 2013). However, it is not known if rodents can adapt attentional intensity based on task utility.

Global arousal, controlled by neuromodulatory systems that make up the reticular activating system (Aston-Jones & Cohen, 2005; McGinley, Vinck, et al., 2015), is thought to be a key physiological mediator of the effects of motivation on attentional intensity. However, three prominent models make widely divergent predictions for the role of arousal in attention. Some evidence, consistent with traditional metabolic and capacity theories, argues that increased arousal is associated with increased attentional performance (Hasselmo & McGaughy, 2004; Massar et al., 2016; Parikh et al., 2007; Saderi et al., 2021). Other evidence suggests just the opposite, that high arousal is associated with task disengagement, and thus that optimal attention occurs with lower arousal (Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011; Usher et al., 1999). A third view is that optimal performance might occur at moderate, rather than high or low levels of arousal (Brink et al., 2016; Faller et al., 2019; McGinley, David, et al., 2015; Schriver et al., 2018; Yerkes & Dodson, 1908). For example, the locus coeruleus noradrenaline system (LC-NA) has been proposed to control the tradeoff between exploitation and exploration (Aston-Jones & Cohen, 2005; Jepma & Nieuwenhuis, 2011), with optimal exploitation (i.e. task-related attention) occurring at moderate levels of tonic firing. A comprehensive approach to behavior and arousal measurement, that could arbitrate between these models, has been lacking.

We previously showed that optimal behavioral and neural detection of simple sounds occur at intermediate levels of pupil size (McGinley, David, 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 major 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), acetylcholine (de Gee et al., 2017; Mridha et al., 2021; Reimer et al., 2016), and perhaps also serotonin (Cazettes et al., 2021) and dopamine (de Gee et al., 2017). Building on these results, we here sought to test whether optimal feature-based attention also exhibits an inverted-U dependence on arousal and whether animals exert self-control on their arousal state to adapt feature-based attention to its utility. We developed an "attentional intensity task" for head-fixed mice and report behavioral and pupillary signatures of strategic attentional intensity allocation in a large cohort (N=88 mice; n=1983 total sessions). The task manipulates coherent motion in acoustic time-frequency space, analogous to random-dot motion used extensively in the visual system (Newsome et al., 1989). We find that during periods of high task utility, mice exhibit multiple signatures of heightened attentional intensity, which is aided by stabilization of arousal closer to an optimal mid-sized level and a reduction in exploratory locomotor behavior.

RESULTS

A feature-based attentional intensity task for head-fixed mice

To study the adaptive recruitment of attentional intensity, and the role therein of global arousal, we developed a feature-based attentional intensity task for head-fixed mice (**Fig. 1**). A large cohort of mice (N=88; n=1983 sessions) were trained to detect coherent time-frequency acoustic motion (called temporal coherence; (Shamma et al., 2011)) embedded, at unpredictable times, in an ongoing random tone-cloud (**Fig. 1A**). The task requires continuous listening and sustained attention to achieve high detection performance. Mice were motivated to listen by being food scheduled and receiving sugar-water reward upon licking during the temporal coherence signal.

We designed the attentional intensity task to be perceptually difficult and therefore highly demanding of attentional resources, so that it would be impossible for mice to sustain consistently high levels of attention and performance across the full session duration (~75 minutes). We manipulated task utility by shifting the reward size back and forth between high (12 ul) and low (2 ul) values, in up to 6 consecutive blocks of 60 trials within each experimental session. Thus, mice should increase attentional intensity during blocks of high reward, to exploit the high utility, and reduce attentional intensity in low-reward blocks. To suppress excessive licking, they received a 14-second time out if and when they licked during the pre-signal tone cloud. We simultaneously recorded pupil size and walking speed as measures of arousal and exploratory behavior, respectively (**Fig 1B**). See Methods for details.

We defined four simple measures of behavioral performance that are well-suited to quantify behavior in this quasi-continuous listening task: (i) Overall response rate, which is the number of responses per second of either noise or signal sounds. Overall response rate captures the overall level of engagement and, in this go/no-go detection task, and is analogous to decision bias (also called criterion in signal detection theory). (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' in signal detection theory). We interpret block-based shifts in discriminatory response rate as the behavioral signature of adaptive allocation of attentional intensity. (iii) Reward rate, which is the fraction of trials that ended in a hit (a response during the signal, which was followed by a sugar 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, the delay from signal onset to correct response. Reaction time is affected by both the efficiency of sensory evidence accumulation and overall tendency to respond, and thus is expected to correlate with both the overall and discriminatory response rates. See Methods for details.

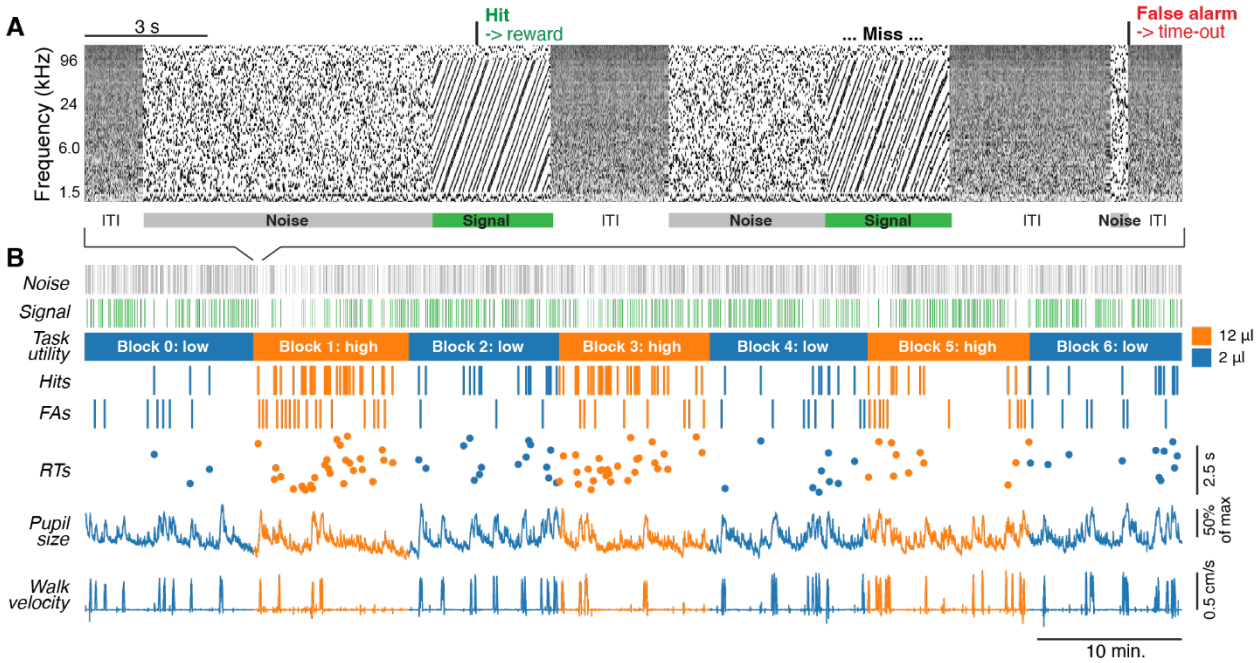

Figure 1. Monitoring behavior and pupil-indexed arousal during the attentional intensity task. (A) Spectrogram of sounds played during three example trials. Mice were head-fixed on a wheel and learned to lick for sugar water reward to report detection of the signal stimulus. Correct go responses (hits) were followed by 2 or 12 mL of sugar water. 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. **(B)** Example session. From top to bottom: noise stimuli, signal stimuli, reward context, correct responses (hits), incorrect responses (false alarms), reaction time (RT) on hit trials, pupil size and walking velocity.

Mice learn to adaptively allocate attentional intensity

Since the adaptive allocation of attentional intensity is a high-level cognitive function, we reasoned that mice may require extensive experience with the value structure of an environment to implement it. To test this, we evaluated the time course of learning for several aspects of the behavior.

The training regime consisted of three phases. 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 µL) 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. 2A**, left). Discriminatory response rate and reward rate also quickly increased across the first few sessions (**Fig. 2D,G,** left). RT gradually decreased by a factor of two during phase 1 (**Fig. S2E**), a further sign that the mice rapidly learned the task structure. In phase 2, we introduced the block-based shifts in reward magnitude and trained the mice until they reached a performance threshold (see Methods). Overall response rate was ~0.3 responses/s at the end of phase 2 (**Fig. 2A**). Through simulations we found that an overall response rate of 0.3 responses/s resulted in the highest reward rate and thus the highest number of rewards (**Fig. S2R**;

Methods). Discriminatory response rate and reward rate also gradually increased (**Fig. 2D,G**), and RT decreased during phase 2 (**Fig. S2E**).

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. 2B**). 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. 2E**). 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. 2E,H**). We fitted exponential functions quantify their 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 (**Fig. S2K,N**). 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 (**Fig. S2D, L,O**).

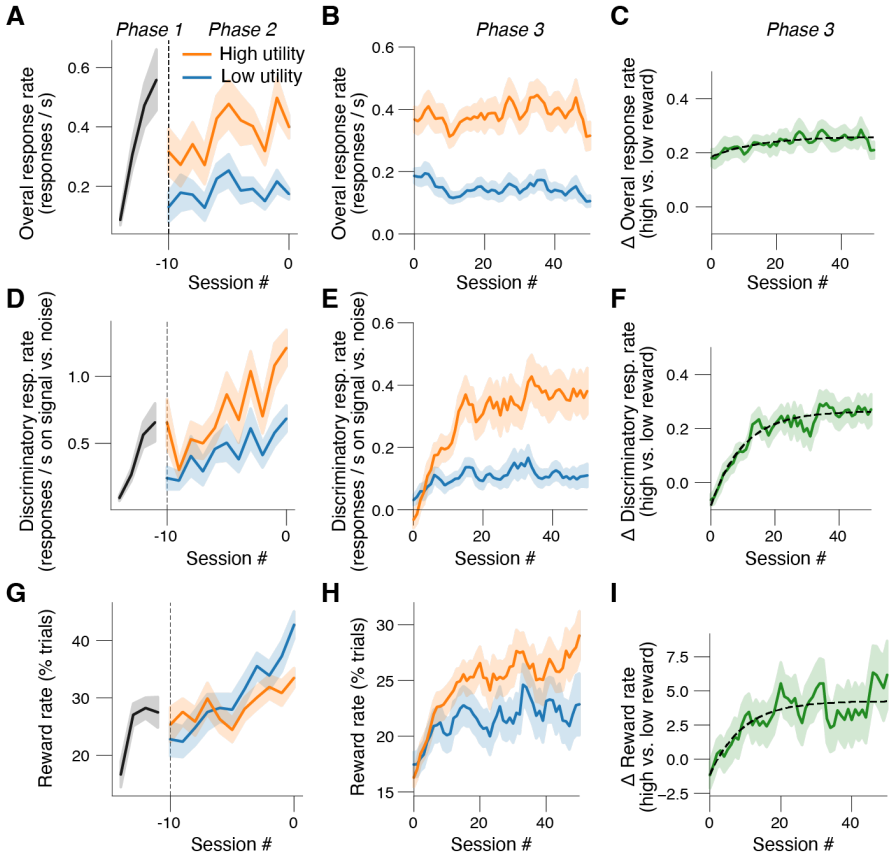

Figure 2. Mice rapidly learn task structure but slowly learn feature-based attention and its allocation. (A) Overall response rate (Methods) across experimental sessions in learning phases 1 and 2; session numbers are with respect to the last session in phase 2. **(B)** Overall response rate across experimental sessions in learning phase 3, separately for high and low reward blocks. **(C)** As B, but for the difference between high and low reward blocks. Dashed line, exponential fit (Methods). **(D-F)** As A-C, but for discriminatory response rate (Methods). **(G-I)** As A-C, but for reward rate (Methods). All panels: shading, 68% confidence interval across animals (N=88, n=1983 sessions).

We also analyzed if and how rapidly mice learn to strategically regulate attentional intensity. 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. 2B,E,H**, orange vs. blue). Reaction times (RTs) were also lower (**Fig. S2F**). 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 (**Fig. 2F,I**), 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.

Performance in a psychometric task variant demonstrates feature-based attention

Due to the temporal structure within each trial, it is theoretically possible that mice could learn to match their reaction times to the temporal statistics of signal occurrence, rather than attend to the temporal coherence in the sound itself. To rule out such a purely 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 fraction of tones that moved coherently through time-frequency space (**Fig. 3A**; 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 temporal anticipation, both discriminatory response rate and reward rate lawfully depended on signal coherence (**Fig. 3B,C**). Thus, mice attended to tone cloud temporal coherence, a complex, high-order auditory feature.

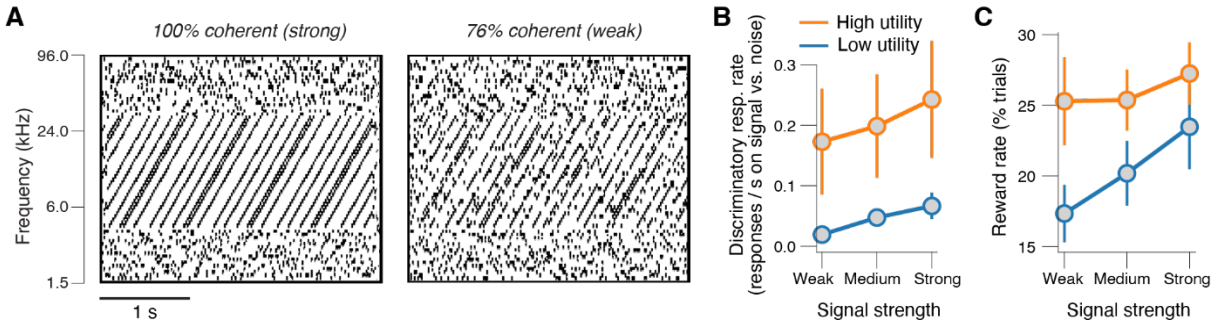

Figure 3. Performance in a psychometric task variant demonstrates feature-based attention. (A) Example spectrograms of strong (left) and weak (right) signals (Methods). **(B)** Discriminatory response rate (Methods) across three difficulty bins and separately for high and low reward blocks. Stats, 2-way 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). **(C)** As D, 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: $F_{2,18} = 2.0$, $p = 0.162$.

Intentional intensity increases rapidly after increases in task utility

We next sought to determine the time scales on which mice adapted their behavior to shifts in task utility, both within and across blocks. We observed that mice spent most of the 1st block (termed block '0'; low reward) becoming engaged in the task (**Fig. S4A-D;** see also **Fig. S7F**). Therefore, we focused all analyses on the subsequent six blocks, between which the available reward size alternated between 12 and 2 μL. See Methods for details.

Across blocks, we found that the animal's overall response rate was higher in the high compared to low reward blocks, across the session (**Fig. 4A-C**). Overall response rate was close to the optimal level (**Fig. S2R**; Methods) in the high reward blocks, while it was substantially lower than optimal in the low reward blocks (**Fig. 4C**). This pattern represents a performance, but not attention, optimization. Crucially, discriminatory response rate was also higher in the high compared to low reward blocks, particularly early in each session (**Fig. 4D-F**). Both overall and discriminatory response rate declined across the session (**Fig. 4A,C,D,F**), likely as a result of fatigue and/or satiety. In sum, an increase in task utility boosts feature-based attentional intensity (discriminatory response rate, analogous to sensitivity) as well as boosting the signal-independent probability of responding (overall response rate, analogous to criterion) to optimal levels.

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 had occurred. Therefore, we analyzed time-dependencies within each block in two different ways: (i) aligned to the switch in (unobserved) reward availability (dashed lines in **Fig. 4A,D,G**), and (ii) aligned to the first hit trial after block transitions (solid lines). After the first high reward in the high reward block, overall response rate immediately increased by 232 ± 8% compared to just before the switch, and after that did not increase further (**Fig. 4B**). 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). After the first high reward in the high reward block, discriminatory response 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. 4E**). When switching from high to low rewards, discriminatory response rate immediately increased by 16 ± 7 % 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). Mice updated their behavior faster when switching from low to high reward, than the other way around. This hysteresis indicates a heightened urgency when task utility is detected to be high.

A rational animal or human should allocate resources in a way that maximizes the reward rate. We thus wondered if mice collected more rewards after increases in task utility. This is not trivially so; since noise stimuli always precede signals in our task, an overall response rate that is too high leads to more false alarms than hits, and thus lowers the expected reward rate (**Fig. S2R**).

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. 4G-I**). An individual difference analysis supported this pattern, in that reward rate was strongly predicted by discriminatory response rate, but not overall response rate, at the per-mouse level (**Fig. S4F**). Mice sustained a stable reward rate of 27.6 ± 0.6 %. across the entire session during the high reward blocks (odd numbered blocks), while reward rate was lower in the low reward blocks from the start (block 2) and declined further across the session (blocks 4 and 6; **Fig. 4G,I**). Mirroring reward rate, RTs were substantially lower in high reward blocks than in low reward blocks, and gradually increased as the experimental session progressed (**Fig. S4H-J**). Reward rate and RT exhibited similar hysteresis effects as the overall and discriminatory response rates (**Figs. 4H, S4I**). Specifically, mice updated their behavior faster when switching from low to high reward than when switching in the other direction.

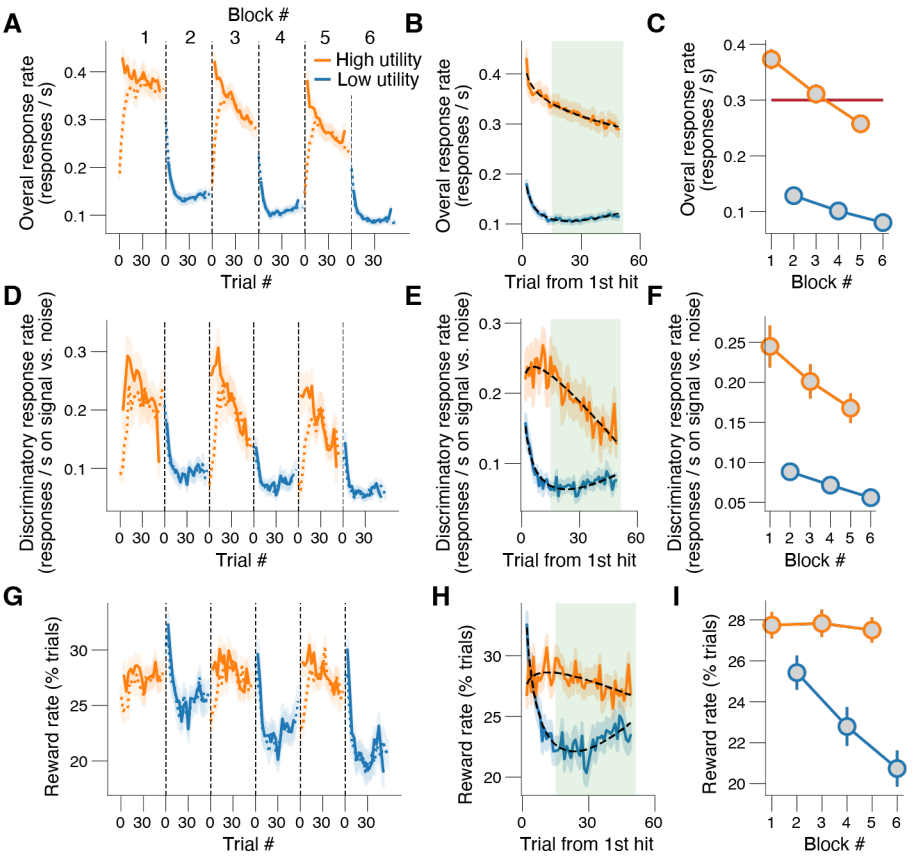

Figure 4. Rapid and adaptive shifts in mouse decision-making after changes in task utility. (A) Overall response rate (Methods) across three high- and three low reward blocks in the same experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit in block. **(B)** As A, but collapsed across blocks of same reward magnitude. Green window, trials used when pooling data across trials within a block (as in panel C). **(C)** As A, 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: $F_{1,87} = 317.3$, $p < 0.001$; 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, n=1983 sessions).

We wondered if the mice who were best at discriminating signals from noise, were also best at allocating attentional intensity, or if attentional allocation relied on a distinctly and variably learned cognitive process. We observed substantial variation between mice in how well they performed the attentional intensity task. Discriminatory response rate (irrespective of task utility) ranged from -0.04 to 0.51, and the change in discriminatory response rate (high- vs low reward blocks) ranged from -0.14 to 0.68. Interestingly, mice who performed best overall were also the best at allocating attentional intensity (**Fig. S4J**). The overall response rate predicted attentional intensity allocation to a substantially lesser extent (**Fig. S4J**), and even negatively so when accounting for the discriminatory response rate (see legend of **Fig. S4J**). Thus, attention and allocation tended to be improved together, but distinctly from response criterion.

For all behavioral metrics, we verified that our results are robust to specifics regarding trial selection (**Fig. S4K-N;** Methods) and are not confounded by effects of time-on-task (**Fig. S4O-R**; Methods). We conclude that mice increase their attentional intensity after increases in task utility: they are more sensitive in discriminating signals from noise, are faster doing so, and they collect more rewards.

Multiple aspects of sensory signal engagement improve during high task utility

We next sought to dissect the elements of the decision-making process underlying the observed effects of task utility on overt behavior (**Figs. 1-4**). Because the target stimulus in our task was a higher-order statistic in ongoing noise, correct detection relies on accumulation across time of partial evidence. Consistent with this interpretation, reaction times were long and variable (see Fig. S2). We therefore applied sequential sampling modeling, as is commonly used for similar tasks in the primate visual system (Gold & Shadlen, 2007). It has been shown that rodents can perform evidence accumulation (Brunton et al., 2013; Hanks et al., 2015). It is unknown if and how evidence accumulation is shaped by task utility.

The widely used drift diffusion model describes the perfect 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 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), this model produces the fastest decisions for a fixed error rate (Bogacz et al., 2006). However, in our task, like most perceptual decisions in nature, the relevant evidence is not stationary. In this case, perfect 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 (**Fig. 5A,B**; (Usher & McClelland, 2001); Methods).

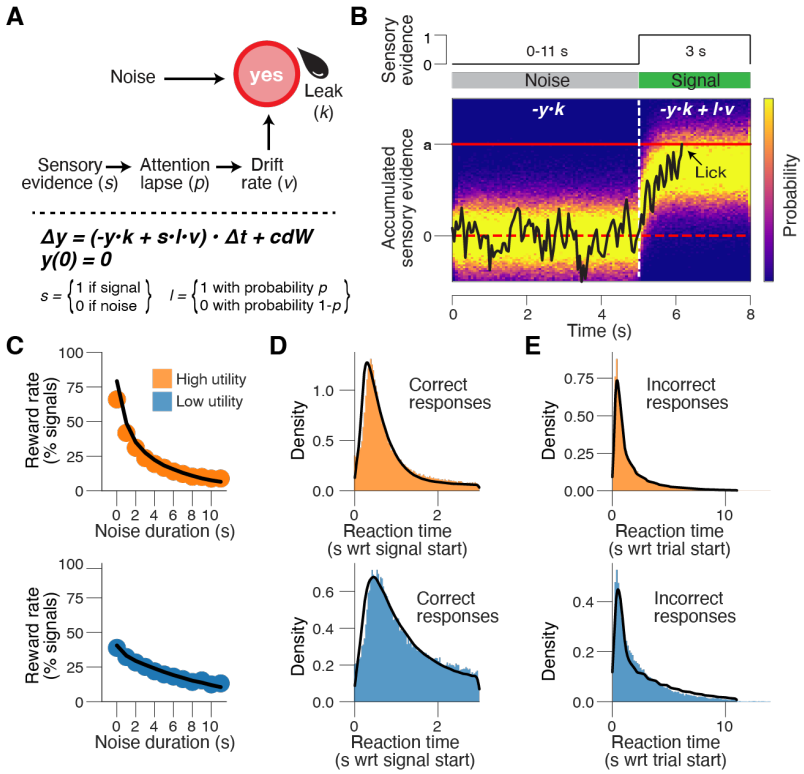

Figure 5. Accumulation-to-bound model accurately and precisely describes decision-making in the attentional intensity task. Schematic of bounded accumulation model accounting for the fraction of go-responses and their associated reaction times (RT). **(B)** The decision-dynamics were governed by leak (k) and gaussian noise during noise stimuli, and additionally by the attention lapse probability (p) and drift rate (v) during the signal stimuli. **(C)** 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. **(D)** RT distribution for correct responses (hits) in the high (top) and low (bottom) reward blocks. Black line, model fit. **(E)** As D, but for incorrect responses (false alarms). Panels C-E: pooled data across animals (N=88, n=1983 sessions).

Our model contained six main parameters (**Fig. 5A,B**; Methods): These 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 (coherence); (iii) attention lapse probability, which is the fraction of signals on which the relevant sensory evidence is not accumulated (together with the drift rate accounted for the fast but infrequent correct responses; **Fig. S5E,F**); (iv) bound height, which controls the speed-accuracy tradeoff and overall response rate; (v) non-decision time, which is the speed of pre-decisional evidence encoding and post-decisional translation of choice into motor response; and (vi) a mixture rate, which sets a fraction of trials on which the mouse makes an automatic response to trial onset, and thus does not engage in evidence accumulation (accounted for the numerous fast incorrect responses; **Fig. S5G,H**). The fitted model accounted well for the behavior in the attentional intensity task, making accurate predictions for choice-outcome probabilities and RTs (**Fig. 5C-E**). We considered three alternative

models (Methods), which provided worse fits, both qualitatively (**Fig. S5E-J**) as well as quantitatively (**Fig. S5K**).

Because of the prominent time-on-task effects we observed in our performance measures (**Fig. 4**), we allowed all model parameters to vary with 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. 6A-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. S6A-C**). 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. S6C**; Methods). There was no significant effect of task utility on the other, non-attention related, model parameters (**Fig. 6D-F**). The biggest effect of time-on-task was on the attention lapse rate (**Fig. 6C**). 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).

The parameter dependencies of leak, drift rate, and attentional lapse rate on task utility all support the conclusion that attentional intensity is greater in high reward blocks. Lower leak indicates more sensory stimulus engagement, which is a signature of attentional intensity. However, in our attentional intensity task, with non-stationary evidence, low leak is not necessarily optimal because more noise is accumulated. Therefore, we used simulations to identify the optimal leak for the task (**Fig. S6J,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 indicates that they were more engaged with the sound stimulus, but also meant that extensive pre-signal noise was integrated, leading to an excess of false alarms and a high overall response 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 (**Fig. S5B**). 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. S6D-I**). In sum, we found that that an increase in task utility resulted in robust changes of the decision computation, at the block-to-block and animal-to-animal levels, all of which support improved attention.

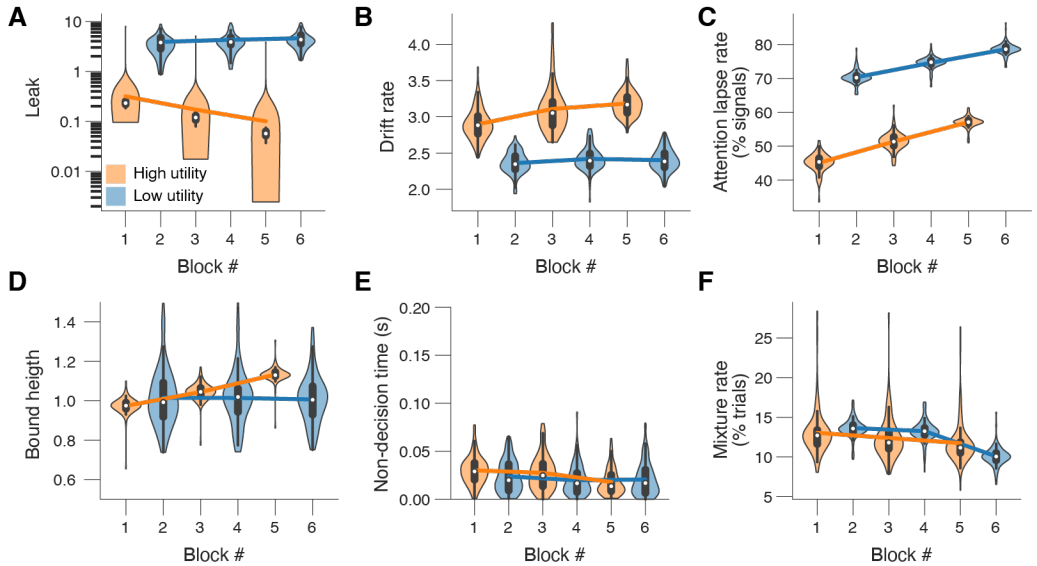

Figure 6. High task utility improves multiple aspects of the decision computation. (A) Fitted leak estimates (kernel density estimate of 100 bootstrapped replicates) separately per block number. Main effect task utility (fraction of bootstrapped parameter estimates in the low reward blocks higher than in the high reward blocks): $p < 0.01$. **(B)** As A, but for drift rate. Main effect task utility: $p < 0.01$. **(C)** As A, but for attention lapse probability. Main effect task utility: $p < 0.01$. Main effect time-on-task (fraction of bootstrapped parameter estimates in the first two blocks higher than in the last two blocks): $p < 0.01$. **(D-F)** As A, but for bound height, non-decision time and mixture rate, respectively. All panels: all further main effects of task utility, time-on-task or interaction effects were not significant ($p > 0.05$).

Optimal performance in the attentional intensity task occurs at intermediate levels of arousal

We previously showed that optimal signal detection behavior in a simple tone-in-noise detection task occurred at intermediate levels of arousal (McGinley, David, et al., 2015). We wondered if a similar inverted-U dependence of performance on arousal was observed in the more-complex attentional intensity task, and in our substantially larger sample size, which likely sampled a broader range of states.

We quantified arousal as the diameter of the pupil measured immediately before each trial (Methods), and first characterized the relationship between pre-trial pupil-linked arousal and behavioral performance, irrespective of the reward context. Consistent with our earlier work, we observed the highest discriminatory response rate and shortest RTs on trials characterized by mid-size pre-trial pupil size (**Fig. 7A,B**; Methods). We additionally observed highest overall response rate on trials with mid-size pupil (**Fig. 7C**), and a linear, negative relationship between reward rate and pre-trial pupil size (**Fig. 7D**). The lack of inverted-U for reward rate occurs because 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, which 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). Therefore, we additionally analyzed behavioral performance as a

function of whether the mice were stationary or locomoting on the wheel. Overall response rate was substantially higher on trials associated with pre-trial walking. However, discriminatory response rate and reward rate were substantially lower and near zero (asterisks in **Fig. 7A,C,D**). These results suggests that responses during walking were essentially random, and thus that attentional intensity was very low.

We finally sought to characterize which computational elements are shaped by the animal's (pupil-linked) arousal state. To that end, we refitted the full model, but now separately for four arousal-defined bins: three pupil-defined bins during stillness, and a separate bin for walking trials (colored background in **Fig. 7A**; Methods). We found a U-shaped dependence of the attention lapse rate on arousal (**Fig. 7G**). Thus, attention to the relevant feature is optimal at intermediate levels of pupil-linked arousal, and worst during walking. Additionally, we observed a general increase of automatic responses with increasing arousal states (**Fig. 7J**). In sum, optimal signal detection behavior in the attentional intensity task occurred at intermediate levels of arousal.

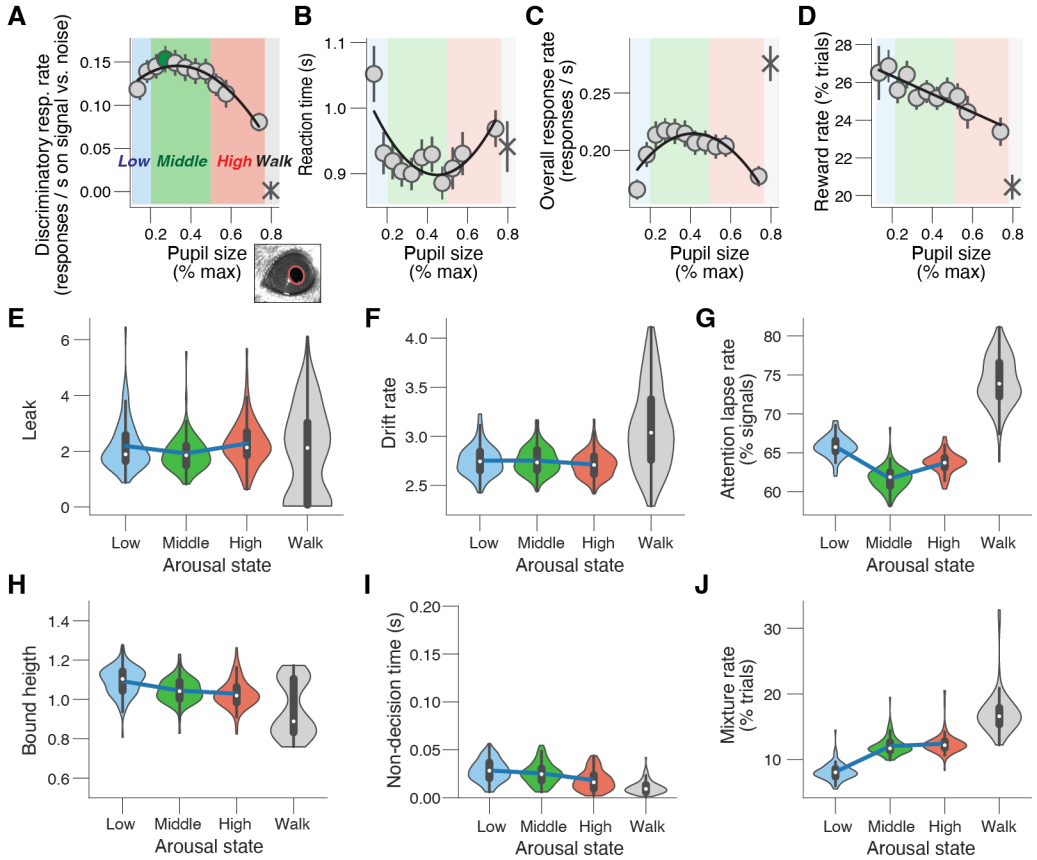

Figure 7. Mice regulate arousal state closer to optimal level during high task utility. (A) Relationship between pre-trial pupil size and discriminatory response rate (irrespective of task utility; Methods). A 1st order (linear) fit was superior to a constant fit ($F_{1,163} = 5.1$, $p = 0.024$) and 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);

error bars, 68% confidence interval across animals (N=88, n=1983 sessions); green data point, optimal pre-trial pupil size bin (Methods); colored panels, binning used for behavioral modeling (panels E-J). **(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)** As A, but for reward rate (Methods). 1st order fit: $F_{1,163} = 5.5$, $p = 0.020$; higher order models were not superior. **(E)** Fitted leak estimates (kernel density estimate of 100 bootstrapped replicates) separately per arousal state (pupil size defined bins corresponding to colors in panels A-D; irrespective of task utility; Methods). No significant differences between arousal states. **(F-J)** As E, but for remaining parameters. The attentional lapse rate was significantly higher in the low vs middle arousal state ($p = 0.05$) and during walking vs the high arousal state ($p < 0.01$). Likewise, the mixture rate was significantly higher in the low vs middle arousal state ($p = 0.01$) and during walking vs the high arousal state ($p = 0.05$). All other comparisons were not significant.

Adaptive shifts in pupil-linked arousal partially mediate improved behavior during high task utility

We finally sought to address whether mice were adaptively regulating their arousal based on task utility, which would result in spending more time close to this optimal state during periods of high task utility. We defined the optimal level of arousal as the pre-trial baseline pupil size bin with maximal discriminatory response rate (green dot in **Fig. 7A**). 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, and less time in the suboptimal low and high arousal states (**Fig. 8A,B**). In line with this, the pre-trial pupil size was also more stable (less variable) in the high vs. low reward blocks (**Fig. 8A** & **Fig. S8L-N**). 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). Mice indeed spent more time close to the optimal arousal state in the high vs low reward blocks (**Fig. 8C**) and were decreasingly effective at regulating their arousal state towards the optimal level as the session progressed (**Fig. 8C**).

We tested whether changes in task utility better predicted changes in the absolute difference between each pre-trial's pupil size and the optimal size (**Fig. 8C**) or in the raw pre-trial pupil size (**Fig. S8F**). First, we 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. Second, we performed a logistic regression of block-wise reward magnitude [0,1] 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 distance from optimal measure (**Fig. S8C**). 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.

Like pre-trial pupil size, walk probability was lower in the high, compared to low, reward blocks (**Fig. S8P-R**). 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 to rest, but instead to engage in alternative aroused and exploratory behaviors, consistent with some models of the function 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 performance effects. To address this, we tested for statistical mediation of arousal in the apparent effect of task utility on the different performance metrics (**Fig. 8D**). 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. 8E**). Similar results of the mediation analysis were observed if we used the pre-trial pupil size, pre-trial pupil size standard deviation or walking probability (**Fig. S8G,O,S**). Taken together, we conclude that regulating pupil-linked arousal towards the optimal level partly implements the adaptive behavioral adjustments to match attention allocation to task utility.

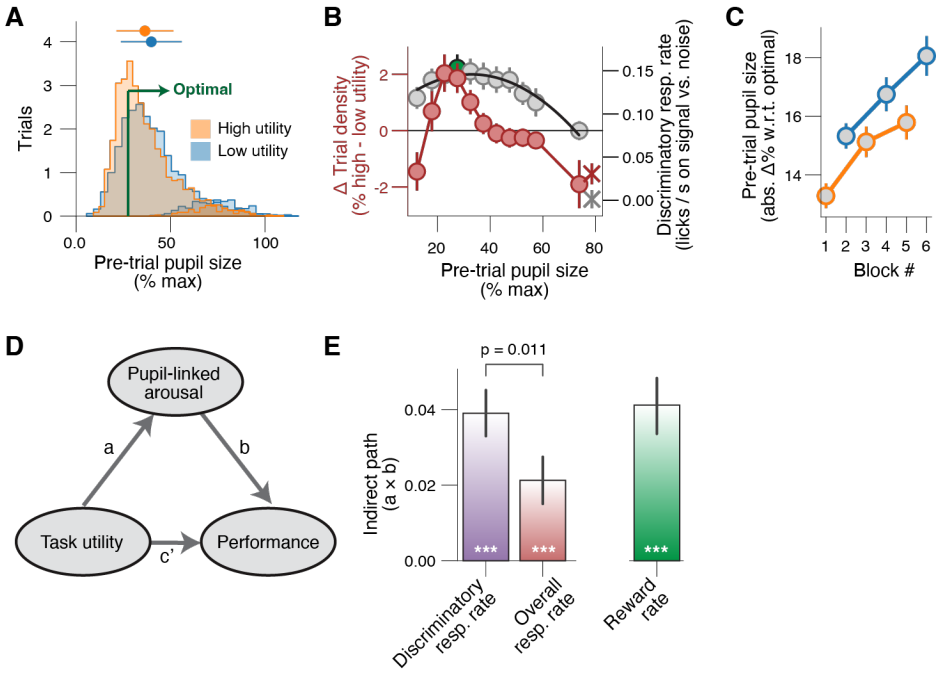

Figure 8. Adaptive changes in pupil-linked arousal partially mediate behavioral improvements seen during high task utility. (A) Histogram of pre-trial pupil measures (Methods) separately for the high and low reward blocks, and for still and walk trials (Methods), from a representative animal. Green line, optimal pre-trial pupil size (Methods). **(B)** In red: the change in trial density after increases in task utility, separately for pupil-defined arousal states; asterisk, walking trials. In grey: same as Fig. 7A (for reference). **(C)** absolute distance from optimal pre-trial pupil size (Methods). Pre-trial pupil size 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: $F_{2,174} = 2.6$, $p = 0.075$. **(D)** 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'. **(E)** Fitted regression coefficients of the the indirect path ($a \times b$; mediation). Stats, Wilcoxon signed-rank test; ***, $p < 0.001$.

DISCUSSION

To efficiently meet their survival needs, organisms must regulate both the selective and intensive aspects of their attention. The study of attention has largely focused on its selective aspect (Carrasco, 2011; Fritz et al., 2007; Maunsell & Treue, 2006), while overlooking the intensive one (but see (Ghosh & Maunsell, 2021)). We here developed a feature-based attention task for head-fixed mice with non-stationary 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. In sum, regulating pupil-linked arousal towards an optimal level partially implements behavioral adjustments that adaptively increase attentional intensity when it is useful.

Our results demonstrate an important, and previously underappreciated, role for motivation in driving attentional intensity. 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 difficulty 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 intensity (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 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 fluctuations in internal state. However, the behavioral functions of this non-stationarity are much less clear. State-dependent neural activity has been observed in primary sensory cortices (Goris et al., 2014; McGinley, David, et al., 2015) and sensory-guided behavior (McGinley, David, et al., 2015) but not in relation to attentional intensity. Recently, spontaneous shifts between engaged, biased, and/or disengaged states have been inferred from behavior (Ashwood et al., 2022; Weilnhammer et al., 2021). Utility-based shifts in behavior (but not attention) have been observed in rodents (Reinagel, 2021; Wang et al., 2013). Other than in humans, the regulation of attentional intensity has only been observed in monkeys (Ghosh &

Maunsell, 2021). 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 (when the reward is large). The adaptive function of low arousal states (such as for online learning and consolidation) and high arousal states (such as for broadly sampling the environment to observe changes and exploring for alternatives), and self-regulation of sampling these states, require further study.

Our results demonstrating motivated shifts in attentional intensity 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 intensity 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).

Our results also demonstrate an important role of arousal in mediating adaptive adjustments in attentional intensity. This role for pupil-indexed arousal 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), which includes 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). Here, motivational factors are customarily neglected; this neglect has been acknowledged but not addressed (Pichora-Fuller et al., 2016). In our attentional intensity task, the perceptual difficulty (temporal coherence) was held constant across the entire session, and 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,' pupil dilation. Future work should determine the interaction of these complementary arousal functions in behavior, perhaps by combining non-stationarity in both task utility and perceptual difficulty.

An important question opened by our results is which neuromodulatory systems contribute to the effects on attentional intensity 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). 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; 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 intensity is more related to discriminatory response rate. Future work is needed to determine the precise roles of neuromodulatory systems in adaptive allocation of attentional intensity.

The specific pattern we observed in the arousal dependence of performance 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 attentional intensity 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 more 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 (but see (de Gee et al., 2020); 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). 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 these 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 disengagement. Thus, 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. S3**) because there was 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

older than 7 weeks at study 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) and ChAT-Cre crossed with Ai162 (N=18; 7 females). 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 experimental manipulations were done 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. 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 consisted of three consecutive intervals (**Fig. 1A**): (i) the "noise" interval, (ii) the "signal" interval, and (iii) the inter-trial-interval (ITI). The duration of the noise interval was randomly drawn beforehand from an exponential distribution with mean of 5 seconds (**Fig. S1A**); 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. S4S-V**), 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" and consisted of consecutive chords of 20 ms duration (gated at start and end with 0.5 ms duration raised cosine). Each chord consisted of 12 pure tones, selected randomly in semitone steps from 1.5-96 kHz. For the signal stimulus, after the semi-random (randomly jittered from tritone log-spaced tones by 1-2 semitones) 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 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. 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.

Training mice to perform the attentional intensity task involved three separate phases. In phase 1, mice performed a version of the task that involved louder signal than noise sounds (58 vs. 52 dB, respectively), several 'classical conditioning' trials (5 automatic rewards during the signal sounds), and no block-based changes in reward magnitude (5 μL after every hit throughout the session). Phase 1 lasted for four experimental sessions (**Fig. S2A**). 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. S2B**). Phase 3 involved the final version of the task, with signal and noise stimuli of equal loudness and without the classical conditioning trials.

In a subset of experiments, the signal quality was systematically degraded by reducing the fraction of tones that moved coherently through 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.

After exclusion criteria (**Fig. S3**), a total of 88 mice performed between 5 and 60 sessions (2100–24,960 trials per subject) of the final version of the attentional intensity task (phase 3), yielding a total of 1983 sessions and 823,019 trials. A total of 10 mice performed the experiment with 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 executed 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. Near-ultraviolet LED light levels were lower during two-photon imaging experiments, to avoid noise on the photo-multiplier tubes.

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. S4A-D;** see also **Fig. S7F**). We found that a fraction of trials began during a lick bout that started in the 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 the 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. S2Q**), 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 time a signal sound was played (in seconds) 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.

To characterize the theoretical relationship between discriminatory response rate and overall response rate (**Fig. S2Q**), and to calculate the optimal overall response rate in our task (**Fig. S2R**), 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 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 phase 3, and a, b and c the free parameters of the fit.

Time course of behavioral adjustments. To calculate the speed of 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. 4B,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 the trial selection procedure (**Fig. S4K-N**).

Accumulation-to-bound modeling. We fitted the 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 the different block numbers (**Fig. 6**) or separately for high and low reward blocks and separately for four arousal-defined bins: three pupil-defined bins during stillness, and a separate bin for walking trials (**Fig. 7E-J**). 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. 5C-E** and **Fig. S5A-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", the decision dynamics were governed by leak and gaussian noise during noise stimuli, and additionally by the drift rate during the signal stimuli (**Fig. 5A,B**):

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

where y is the decision variable (black example trace in **Fig. 5B**), 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 a (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", 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" 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" 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).

To calculate the optimal leak in the attentional intensity task (**Fig. S6J,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. 6**). 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 (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 (McGinley, David, et al., 2015; Mridha et al., 2021).

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. S7A-C**), 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. S8H-K**). 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 dot in **Fig. 7A**).

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. S7D**).

Statistical comparisons

We used a one-way repeated measures ANOVA to test for the effect of signal coherence (**Fig. 3B,C**). 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. 4C,F,I** & **Fig. 8C**).

We directly compared bootstrapped distributions of the model parameter estimates to test for the main effect of task utility and time-on-task (**Fig. 6**) or the effect of arousal state (**Fig. 7E-**

J). 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. S5K**). 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. 7A-D**):

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

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 to characterize the interaction between task utility, arousal, and attention (**Fig. 8D,E**). 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

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), 201–212. 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

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

Brink, R. L. van den, Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil Diameter Tracks Lapses of Attention. *PLOS ONE*, *11*(10), e0165274. https://doi.org/10.1371/journal.pone.0165274

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

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

Faller, J., Cummings, J., Saproo, S., & Sajda, P. (2019). Regulation of arousal via online neurofeedback improves human performance in a demanding sensory-motor task. *Proceedings of the National Academy of Sciences*, *116*(13), 6482–6490. https://doi.org/10.1073/pnas.1817207116

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), 858–865. 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

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*.

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), 1–15. 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.

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.

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), 1539. 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.

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.

Parikh, V., Kozak, R., Martinez, V., & Sarter, M. (2007). Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron*, *56*(1), 141–154.

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

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

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

Roesch, M. R., & Olson, C. R. (2004). Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex. *Science*, *304*(5668), 307–310. https://doi.org/10.1126/science.1093223

Saderi, D., Schwartz, Z. P., Heller, C. R., Pennington, J. R., & David, S. V. (2021). Dissociation of task engagement and arousal effects in auditory cortex and midbrain. *ELife*, *10*, e60153. https://doi.org/10.7554/eLife.60153

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

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience*, *1*(3), 199–207. https://doi.org/10.1038/35044563

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

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*.

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

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

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.

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

END NOTES

Acknowledgements

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

Author contributions

Conceptualization: JWG, MJM
Funding Acquisition: MJM
Data acquisition software: WZ
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

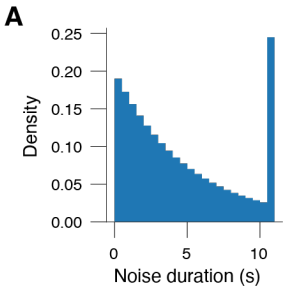

Figure S1. (A) Histogram of noise durations. In most sessions (82.8 %), randomly drawn noise durations greater than 11 s were set to 11 s (resulting in the last bin in the histogram). In the remainder of sessions (17.2 %), these trials were turned into catch trials consisting of 14 seconds of noise.

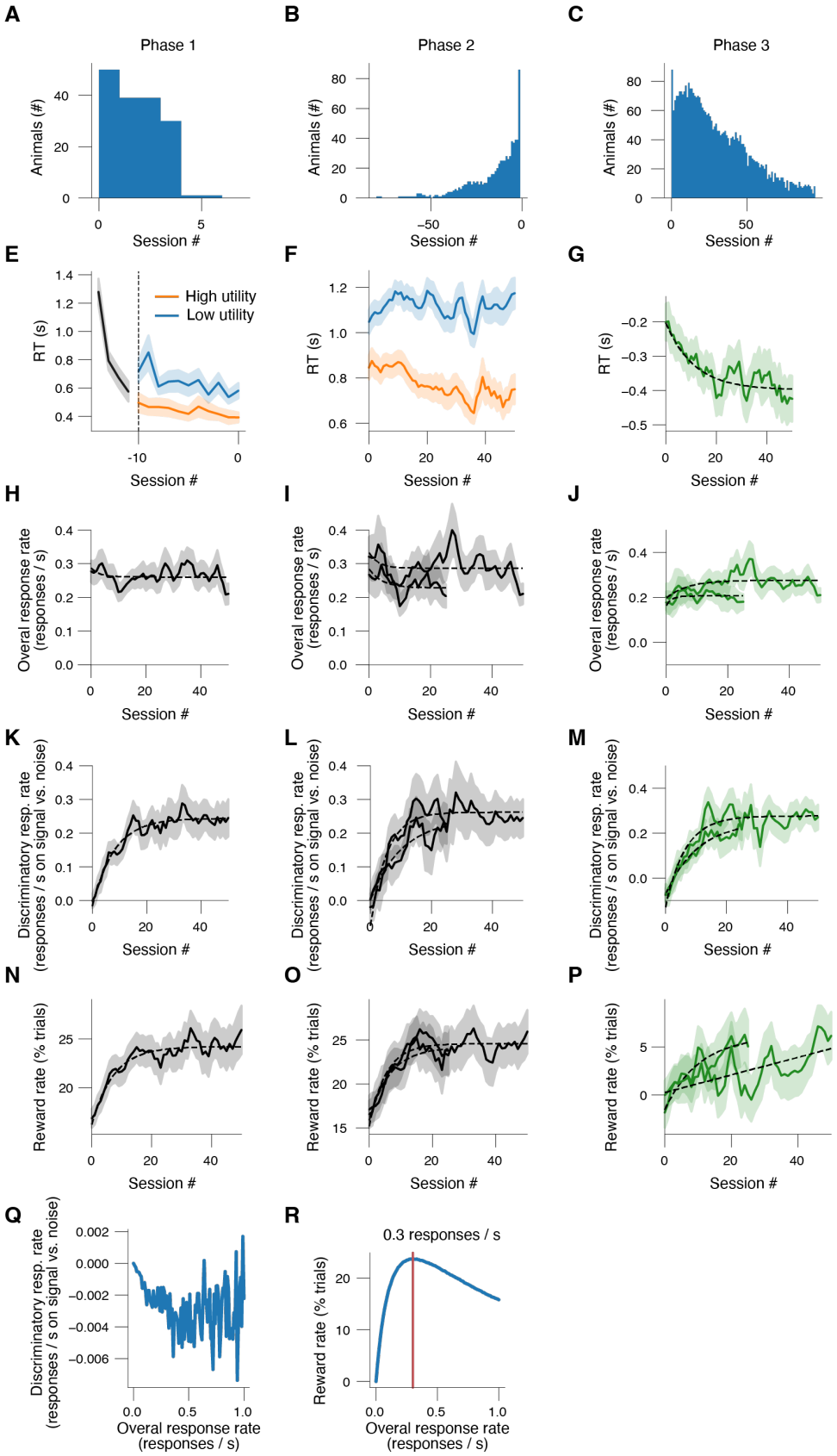
A
Phase 1
Animals (#)
Session #
B
Phase 2
Animals (#)
Session #
C
Phase 3
Animals (#)
Session #
D
Animals (#)
Max session #
E
RT (s)
High utility
Low utility
Session #
F
RT (s)
Session #
G
RT (s)
Session #
H
Overal response rate (responses / s)
Session #
I
Overal response rate (responses / s)
Session #
J
Overal response rate (responses / s)
Session #
K
Discriminatory resp. rate (responses / s on signal vs. noise)
Session #
L
Discriminatory resp. rate (responses / s on signal vs. noise)
Session #
M
Discriminatory resp. rate (responses / s on signal vs. noise)
Session #
N
Reward rate (% trials)
Session #
O
Reward rate (% trials)
Session #
P
Reward rate (% trials)
Session #
Q
Discriminatory resp. rate (responses / s on signal vs. noise)
Overal response rate (responses / s)
R
0.3 responses / s
Reward rate (% trials)
Overal response rate (responses / s)

Figure S2. (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. **(E)** RT on hit trials (Methods) across experimental sessions in learning phases 1 and 2; session numbers are with respect to the last session in phase 2. **(F)** RT on hit trials across experimental sessions in learning phase 3. **(C)** As F, but for the difference between high and low reward blocks. Dashed line, exponential fit (Methods). **(H)**. Overall response rate across experimental sessions in learning phase 3, collapsed across reward context. Dashed line, exponential fit (Methods). **(I)** As H, but separately for mice who performed at least 25 sessions in phase 3, and mice who performed at least 50 sessions in phase 3. **(J)** As I, but for the difference between the high and low reward blocks. **(K-M)** As H-J, but for discriminatory response rate. **(N-P)** As H-J, but for reward rate. **(Q)** Simulation of relationship between discriminatory response rate and overall response rate (Methods). **(R)** Simulation of relationship between reward rate and overall response rate (Methods). Red line, optimal overall response rate (reward rate peaks). Panels E-P: shading, 68% confidence interval across animals (N=88; n=1983 sessions).

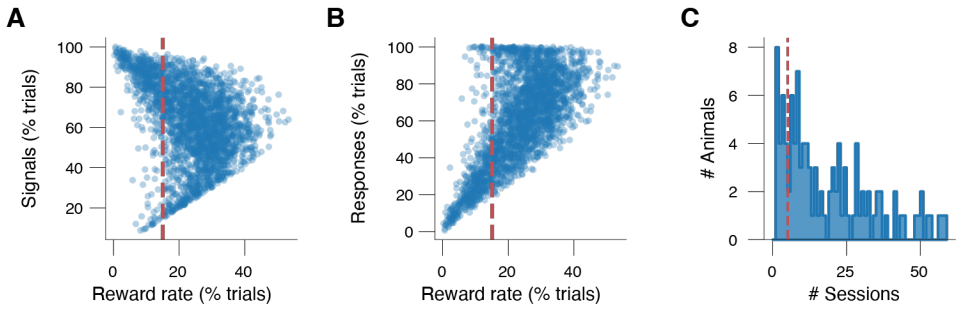

Figure S3. (A) Fraction of trials containing a signal plotted against reward rate (after excluding sessions in panel E). Every data point is a unique session. We excluded 455 sessions with a reward rate smaller than 15%, thereby excluding 4 animals. **(B)** As G, but for fraction of trial on which the animal responded on the y-axis. **(C)** 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.

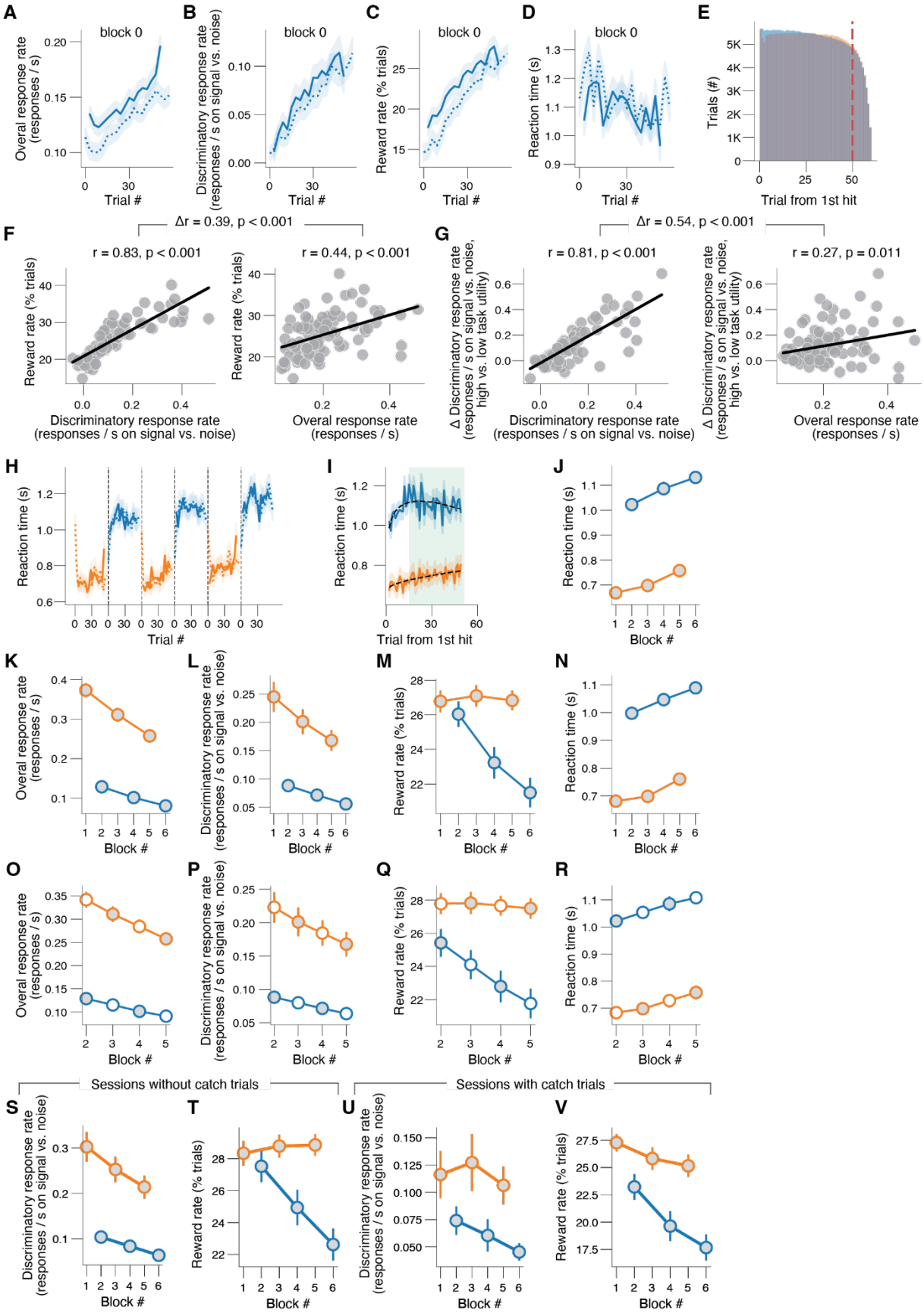
A
block 0
Overal response rate (responses / s)
Trial #
B
block 0
Discriminatory response rate (responses / s on signal vs. noise)
Trial #
C
block 0
Reward rate (% trials)
Trial #
D
block 0
Reaction time (s)
Trial #
E
Trials (#)
Trial from 1st hit
F
Δr = 0.39, p < 0.001
r = 0.83, p < 0.001
Reward rate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

Figure S4. (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)** As A, but for reaction time (RT) on hit trials. **(E)** Histogram of trial number counted from first hit in a block (across all animals and experimental sessions). Dashed red line, cutoff for plotting in panel H and main Fig. 4A,D,G. **(F)** 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). **(G)** 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). **(H)** 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. **(I)** As H, but collapsed across blocks of same reward magnitude. Green window, trials used when pooling data across trials within a block (as in panel J). **(J)** As H, 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. **(K-N)** As panel J and main Fig. 4C,F,I,L, 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: $F_{1, 87} = 638.3$, $p < 0.001$. **(O-R)** As panel J and main Fig. 4C,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: $F_{1, 87} = 411.6$, $p < 0.001$. **(S,T)** As main Fig. 4F,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. **(U,V)** As main Fig. 4F,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).

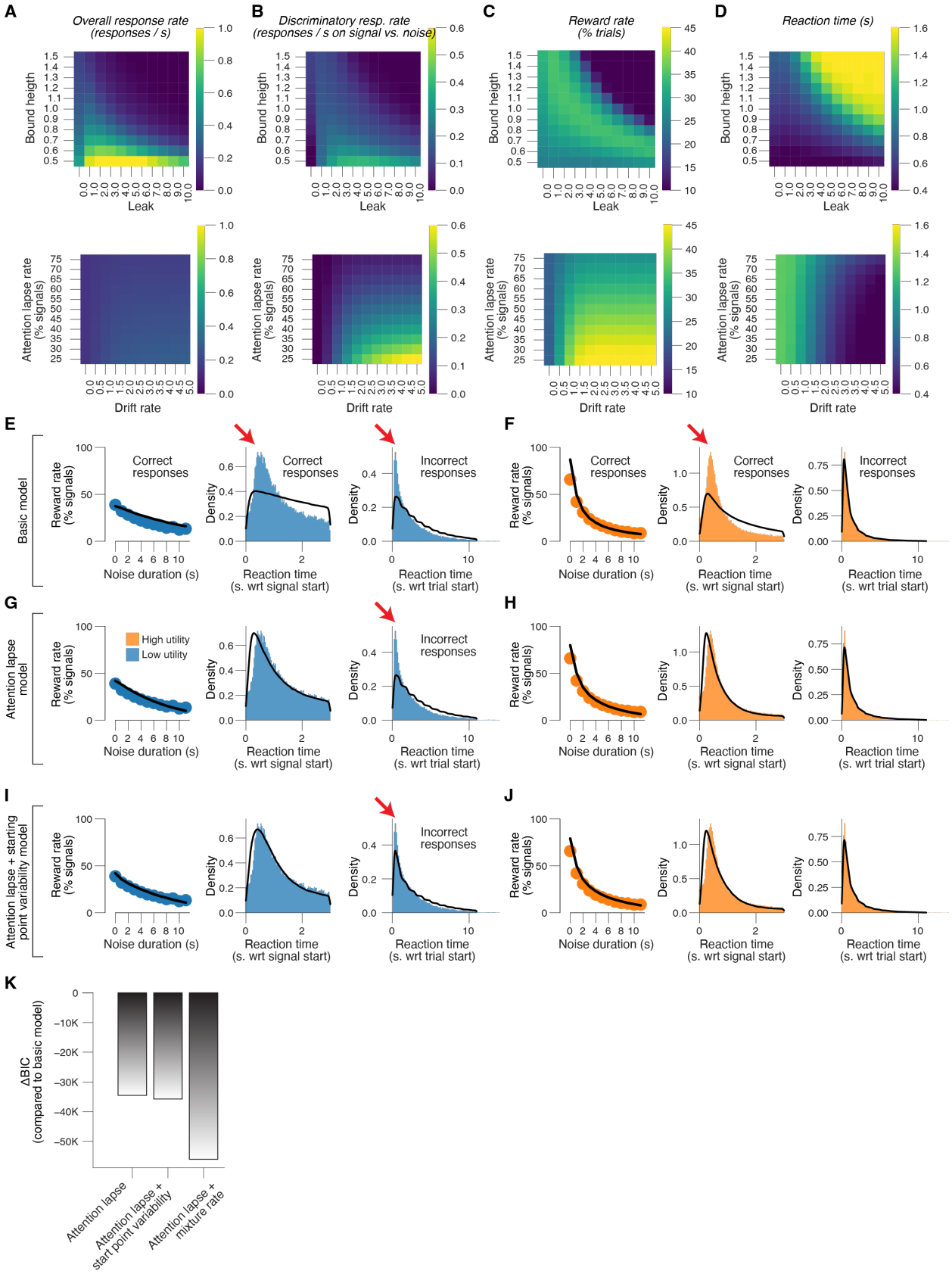
A
Overall response rate (responses / s)
B
Discriminatory resp. rate (responses / s on signal vs. noise)
C
Reward rate (% trials)
D
Reaction time (s)
Bound heigth
Leak
Attention lapse rate (% signals)
Drift rate
E
F
G
H
I
J
Basic model
Attention lapse model
Attention lapse + starting point variability model
Reward rate (% signals)
Noise duration (s)
Correct responses
Incorrect responses
Density
Reaction time (s. wrt signal start)
Reaction time (s. wrt trial start)
High utility
Low utility
K
ΔBIC (compared to basic model)
Attention lapse
Attention lapse + start point variability
Attention lapse + mixture rate

Figure S5. (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.

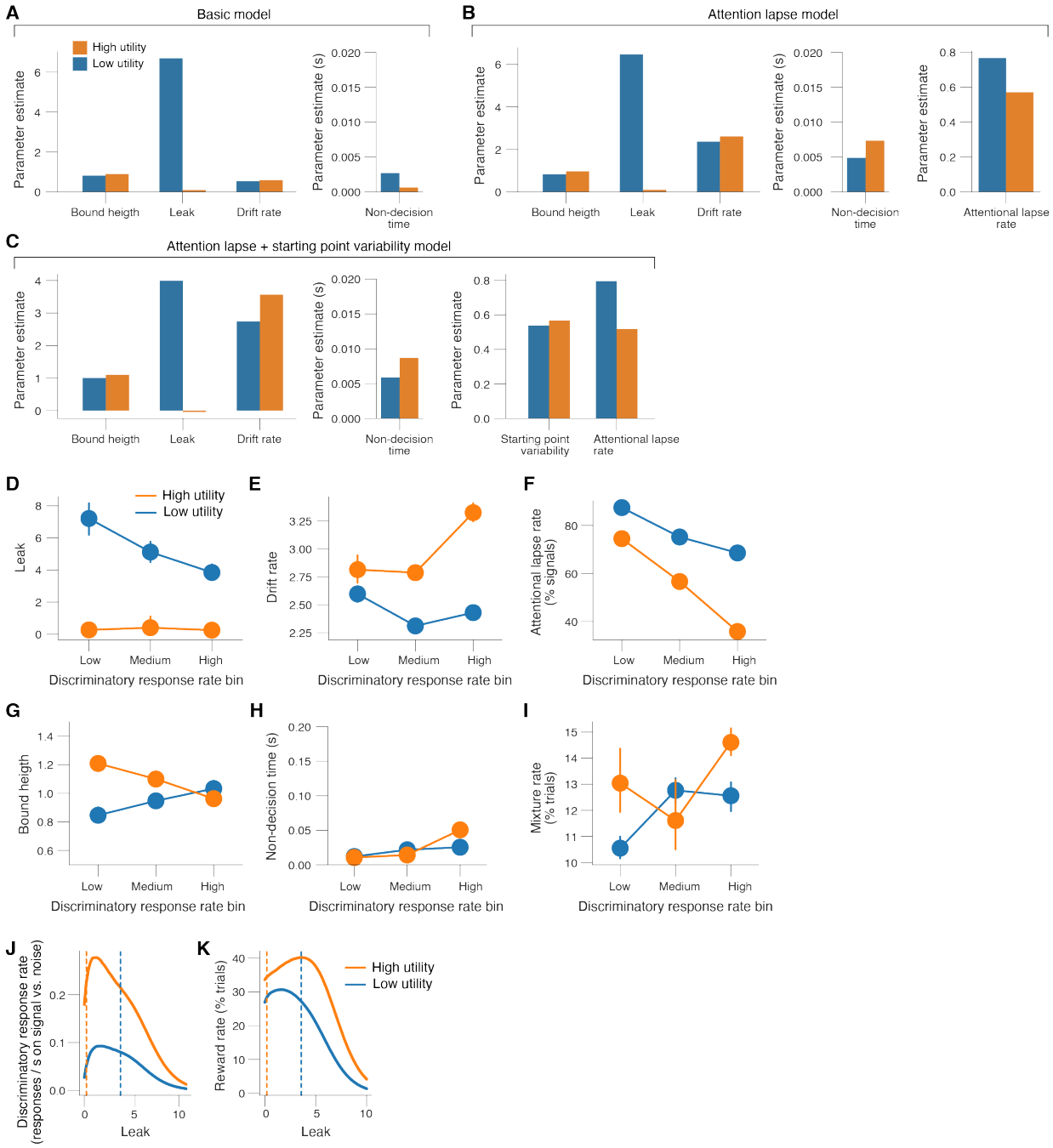

Figure S6. (A) Parameter estimates from the "basic model" (Methods), separately per block type. **(B)** As A, but for the "attention lapse" model (Methods). **(C)** As A, but for the "attention lapse + starting point variability model" (Methods). **(D-I)** Parameter estimates from the "full model" (mean 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 30, 29 and 29 mice with the lowest, middle and highest mean discriminatory response (irrespective of task utility), respectively. **(J)** 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. **(K)** As J, but for simulated reward rate.

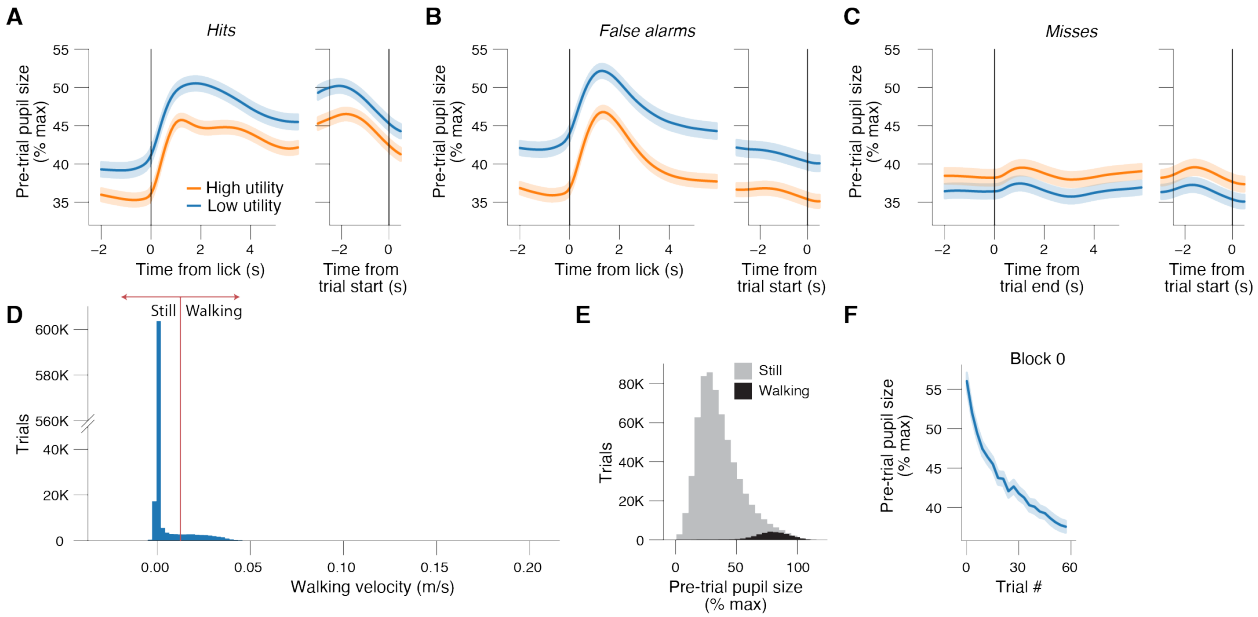

Figure S7. (A) Pupil size across time on hit trials, separately for high and low reward blocks. Left, locked to licks; right, locked to next trial's onset. **(B)** As A, but on false alarm trials. **(C)** As A, but for miss trials. Left, locked to trial offset (end of 3-s signal sound). **(D)** Histogram of pre-trial walking velocity (across all animals and experimental sessions; Methods). Red line, cutoff for defining walking. **(E)** Histogram of pre-trial pupil size (across all animals and experimental sessions), separately for still and walking trials. **(F)** 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. All panels: shading or error bars, 68% confidence interval across animals (N=88; n=1983 sessions).

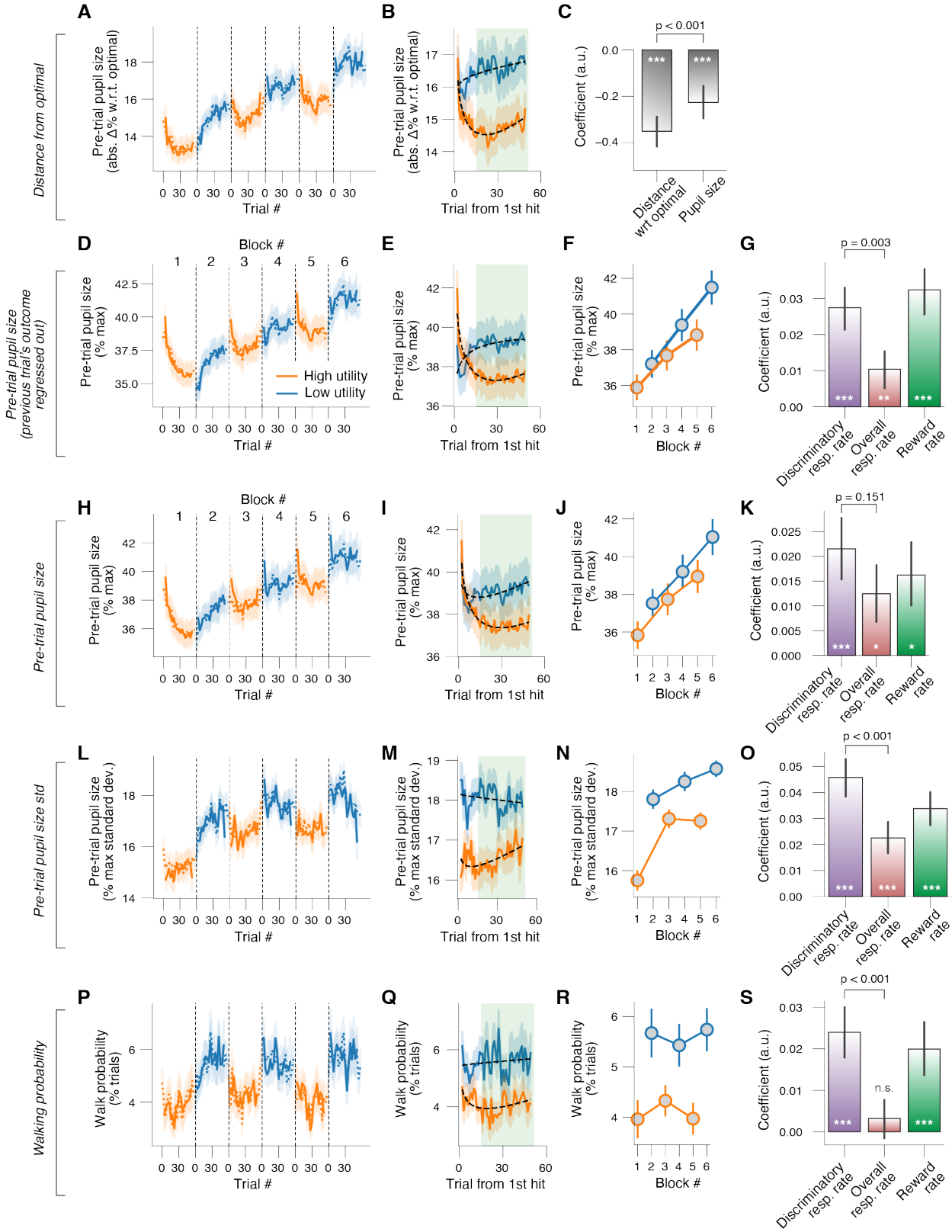

Figure S8. (A) Absolute distance from optimal pre-trial pupil size across three high and three low reward blocks in the same experimental session. Dashed lines, locked to block onset; solid lines, locked to first hit

in block. **(B)** As A, but collapsed across blocks of same reward magnitude. **(C)** Coefficients from logistic regression of block-wise task utility [high vs. low] on either mean pre-trial pupil size (left) or absolute distance from optimal pre-trial pupil size (right). Stats, Wilcoxon signed-rank test; ***, $p<0.001$. **(D)** 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. **(E)** As D, but collapsed across blocks of same reward magnitude. **(F)** As E, 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: $F_{2,174} = 6.3$, $p = 0.002$. **(G)** With pre-trial pupil size as mediation, fitted regression coefficients of the indirect path ($a \times b$; mediation). Stats, Wilcoxon signed-rank test; *, $p < 0.05$; ***, $p < 0.001$. **(H-K)** As D-G, but for pre-trial pupil size without previous trial's outcome regressed out (Methods). Stats in panel J: 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: $F_{2,174} = 1.3$, $p = 0.283$. **(L-O)** As D-G, but for standard deviation of pre-trial pupil size. Stats in panel N: 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: $F_{2,174} = 8.4$, $p < 0.001$. **(P-S)** As D-G, but for walk probability (Methods). Stats in panel R: 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).