Phasic arousal optimizes decision computations in mice and humans

J. W. de Gee[1,2,3,4], K. Tsetsos[1], D. A. McCormick[5,6], M. J. McGinley[3,4,6*,#], T. H. Donner[1,2,7*]

[1]Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [2]Department of Psychology, University of Amsterdam, Amsterdam, Netherlands; [3]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA; [4]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA; [5]Institute of Neuroscience, University of Oregon, OR, USA; [6]Department of Neuroscience, Yale University, New Haven, CT, USA; [7]Amsterdam Brain & Cognition, University of Amsterdam, Amsterdam, Netherlands.

(* shared senior / corresponding authorship; # lead contact)

Neuromodulatory brainstem systems controlling the global arousal state of the brain are phasically recruited during cognitive tasks. The function of such task-evoked neuromodulatory signals is debated. Here, we uncovered a general principle of their function, across species and behavioral tasks: counteracting maladaptive biases in the accumulation of evidence. We exploited that neuromodulatory brainstem responses are mirrored in rapid dilations of the pupil. Task-evoked pupil dilations predicted smaller biases in an auditory detection task performed by humans and mice. In humans, this effect generalized to a more complex, and well-known form of bias: risk-seeking in a value-based ("stock-market") decision. Across tasks, pupil-linked bias suppressions were specifically due to changes in the accumulation of evidence, indicating that phasic changes in arousal state shape the formation of the decision. Thereby, phasic arousal accounts for a significant component of the trial-to-trial variability overt choice behavior.

INTRODUCTION

Even when awake, the central arousal state of the brain changes on a moment-to-moment basis (Aston-Jones and Cohen, 2005; McGinley et al., 2015b). These state changes are controlled in large part by modulatory neurotransmitters released from brainstem nuclei with widespread projections to the rest of the brain, including most parts of the cerebral cortex. For example, the noradrenergic locus coeruleus (LC) and cholinergic basal forebrain have widespread projections to cortex, and change the operating mode of their cortical target circuits, such as the gain of synaptic interactions (Aston-Jones and Cohen, 2005; Harris and Thiele, 2011; S.-H. Lee and Dan, 2012). It is well established that these neuromodulatory systems are closely involved in cognition: they are rapidly recruited during elementary decisions, signaling states such as decision uncertainty or surprise (Aston-Jones and Cohen, 2005; Bouret and Sara, 2005; Dayan and Yu, 2006; Lak et al., 2017; Parikh et al., 2007).

Yet, the functional consequences of task-evoked neuromodulatory responses remains debated. One account holds that task-evoked LC responses during decision-making are triggered by a threshold crossing in some cortical circuit accumulating evidence for different choice options (Aston-Jones and Cohen, 2005). The resulting cortex-wide noradrenaline release then facilitates the translation of the choice into a motor act. This model has been supported by observations from monkeys performing simple perceptual decisions, that phasic responses of LC neurons are more closely aligned to behavioral responses than to the onset of stimuli (Clayton et al., 2004; Rajkowski et al., 2004) and encode decisions to execute, but not withhold, movements (Kalwani et al., 2014).

Another account (Dayan and Yu, 2006) holds that task-evoked neuromodulatory responses occur throughout the decision formation process. In particular protracted decisions (1/2 s or longer) result from the accumulation of fluctuating evidence over time up to a critical decision bound (Bogacz et al., 2006; Brody and Hanks, 2016; Gold and Shadlen, 2007; Ratcliff and McKoon, 2008). Here, changes in central arousal state may alter evidence accumulation and thus choice outcome. This account has recently obtained support from studies of task-evoked pupil responses in humans, which built on the observation that pupil dilation indexes cortical arousal state (Larsen and Waters, 2018; McGinley et al., 2015b) as well as LC responses (Larsen and Waters, 2018) in humans (de Gee et al., 2017; Murphy et al., 2014), monkeys (Joshi et al., 2016; Varazzani et al., 2015), and mice (Liu et al., 2017; Reimer et al., 2016). In line with an intra-decisional account, pupil dilation is driven throughout protracted decision-formation (de Gee et al., 2014; Murphy et al., 2016) and predicts changes in evidence accumulation (Cheadle et al., 2014; de Gee et al., 2017).

It is not known if the discrepancy between the intra- and post-decisional models of phasic arousal is due to differences in species (e.g. between monkeys and humans), physiological signals, or behavioral tasks. Furthermore, it is unclear if the functional consequences of task-evoked neuromodulatory responses generalize to higher-level decisions based on more abstract evidence. Many of the decisions humans care about (e.g. which career to pursue, or which stock to buy) are of the latter kind and more protracted than the elementary perceptual decisions previously used in this field. If phasic arousal indeed interacts with the decision formation, as postulated by the intra-decisional account, more protracted decisions should provide even more room arousal to exert its impact on choice behavior.

Here, we identified a general principle of the function of task-evoked neuromodulatory responses. We combined pupillometry and behavioral modeling in humans and mice performing the same simple perceptual decision (auditory go/no-go detection). In addition, humans performed a forced-choice decision task based on identical stimuli as well as a value-based choice task often used to study stock market decisions. In all both species and all tasks, pupil responses predicted a suppression of a bias in evidence accumulation. Thus, the ongoing deliberation culminating in a choice (Shadlen and Kiani, 2013) is shaped by task-evoked neuromodulatory responses.

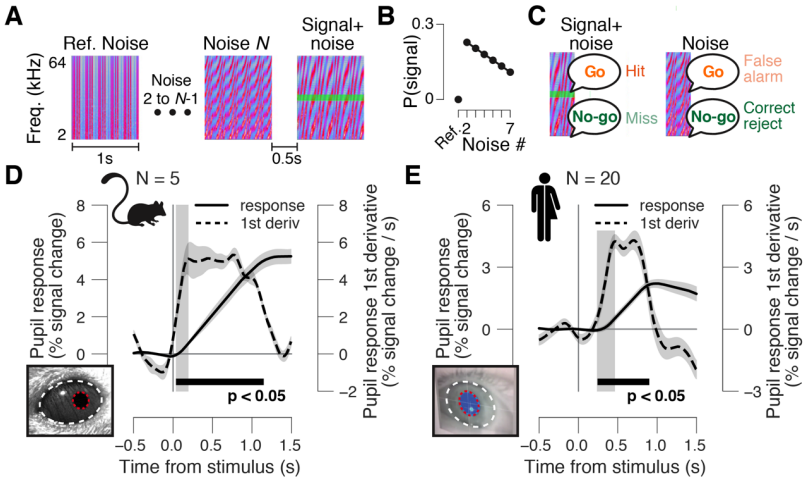

Figure 1. Quantifying task-evoked pupil responses during an elementary decision. (A) Auditory go/no-go tone-in-noise detection task. Schematic sequence of stimuli during a trial. Subjects responded to a weak signal (pure tone) in noise and withheld a response for noise sounds. Each sound was treated as a separate decision (Materials and Methods). **(B)** Probability of target signal across the 2-7 sounds. Hazard rate of signal onset was

approximately flat across the trial. **(C)** Combination of stimulus category (signal+noise vs. noise) and behavioral choice (go vs. no-go) yielded signal detection theory categories. **(D)** Task-evoked pupil response (solid line) and response derivative (dashed line) in mice. Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. **(E)** As D, but for humans. D,E: group average (N = 5; N = 20); shading, s.e.m.; stats, paired-samples t-test.

RESULTS

Phasic arousal predicts reduction of perceptual choice bias in mice and humans

We first trained mice (N = 5) and humans (N = 20) to report detection of a near threshold auditory signal (Fig. 1A,B; Materials and Methods). Because target signals were embedded in fluctuating noise, detection performance could be maximized by accumulating the sensory evidence over time. Subjects searched for a target (pure tone) embedded in a sequence of discrete, but dynamic, noise tokens. To indicate a yes-choice, mice licked for sugar water reward and human subjects pressed a button. The target-signal strength was manipulated by varying the sound level of the tone, while keeping the noise level constant.

To track phasic arousal, we measured the rising slope of the pupil, immediately after sound onset. We choose this measure for three reasons: (i) to most specifically track noradrenergic activity (Reimer et al., 2016), (ii) for its temporal precision in tracking arousal during fast-paced tasks (Fig. 1D,E), and (iii) to eliminate contamination by movements (licks and button-presses) (de Gee et al., 2014; Hupé et al., 2009) (Materials and Methods). Pupil responses occurred within 240 ms after stimulus onset (Fig. 1D,E) and occurred also on trials without a behavioral response (Fig. S1A,C), consistent with other observations (C. R. Lee and Margolis, 2016; Schriver et al., 2018).

In both mice and humans, we found a consistent relationship between the early, task-evoked pupil response and decision outcome, in line with the intra-decisional account of phasic arousal function. We quantified detection performance in terms of the signal detection-theoretic measures (Green and Swets, 1966) sensitivity (d') and bias (criterion), as well as reaction times (RT) (Materials and Methods). We found that both mice and humans had an overall conservative decision criterion (Fig. 2). Optimality analysis of the go/no-go choice task showed that this conservative bias was maladaptive, reducing the fraction of correct/rewarded choices below what could be achieved at a given perceptual sensitivity (Fig. S2). Critically, this maladaptive bias was suppressed on trials with large pupil responses. This was true across signal strengths for both species (Fig. 2). Pupil responses exhibited a less consistent relationship to perceptual sensitivity and RT (Fig. S3).

Previous work has associated baseline, pre-stimulus arousal state with non-monotonic (inverted U-shape) effects on decision performance (Aston-Jones and Cohen, 2005; McGinley et al., 2015a; Yerkes and Dodson, 1908). Here, for both species, the dominant predictive effect of task-evoked pupil responses on criterion was linear (Fig. 2, solid lines). In the mice, there was additionally a small, but significant, non-linear component (dashed lines in Fig. 2A; see Materials and Methods). In sum, pupil-linked phasic arousal was consistently associated with a robust, primarily linear, reduction of maladaptive perceptual bias in mice and humans.

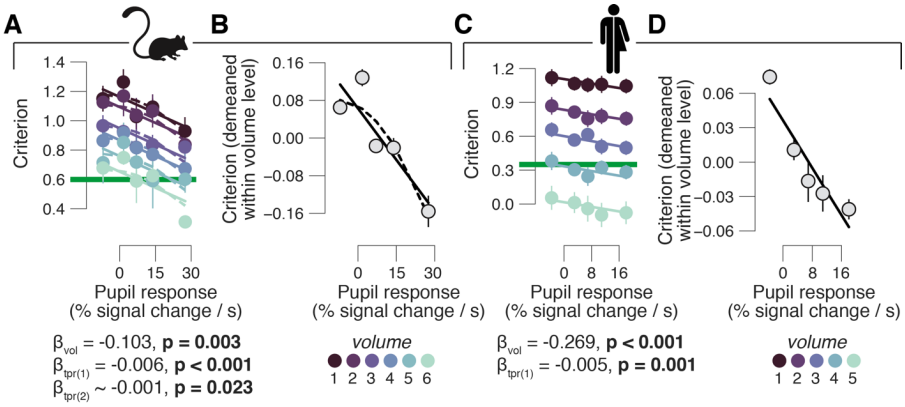

Figure 2. Phasic arousal predicts reduction of perceptual decision bias. (A) Relationship between choice bias (quantified by signal detection criterion) and task-evoked pupil response in mice, separately for different signal strengths. Green lines, optimal criteria (see Fig. S2). Linear fits were plotted if first-order fit was superior to constant fit; quadratic fits (dashed lines) where plotted if second-order fit was superior to first-order fit (Materials and Methods). Stats, mixed linear modeling (see Materials and Methods). **(B)** As A, but after demeaning criterion per signal strength. **(C,D)** As A,B, but for humans. All panels: group average (N = 5; N = 20); error bars, s.e.m.

Phasic arousal predicts a reduction of evidence accumulation bias

We fitted the drift diffusion model (Ratcliff and McKoon, 2008) to gain insight into how choice biases were suppressed in association with pupil responses. The diffusion model belongs to a class of "sequential sampling" models (Bogacz et al., 2006; Brody and Hanks, 2016; Gold and Shadlen, 2007; Ratcliff and McKoon, 2008) that describe the accumulation of noisy sensory evidence in a latent decision variable that drifts to one of two bounds. The diffusion model accounts well for behavioral data from a wide range of two-choice and go/no-go tasks (Ratcliff et al., 2016; Ratcliff and McKoon, 2008). We here used the diffusion model to distinguish between two possible mechanisms that can bring about a change in choice bias: changing the amount of evidence required to reach one over the other decision bound, changing the accumulation process itself (Ratcliff and McKoon, 2008).

Both mechanisms are distinguishable through their distinct effects on the shape of the RT distribution (de Gee et al., 2017; Ratcliff and McKoon, 2008; Urai et al., 2018). Specifically, a shift in starting point towards the "go"-bound yields more go-choices primarily for short RTs (Fig. 3A, left), whereas the effect of a bias in the drift grows with time, yielding more go-choices also for long RTs (Fig. 3A, right). We first compared two models: the "starting point model" (the simplest manifestation of an accumulation bias), in which all parameters except drift criterion were allowed to vary with pupil response; the "drift criterion model", in which all parameters except starting point were varied. In both models, we fixed drift rate variability across the five pupil-bins, because it was not relevant to our hypothesis and is difficult to constrain (Ratcliff and McKoon, 2008; Wiecki et al., 2013).

In both species, we found that the bias suppression associated with a large pupil response was due to a shift in drift criterion (Fig. 3). First, formal model comparison favored a change in drift criterion over a change in starting point (Fig. 3B; Materials and Methods). Second, across species and signal strengths, we found a positive linear relationship between pupil responses and drift criterion (Fig. 3D,E,H,I). The drift criterion model almost fully accounted for the pupil dependent changes in overt decision bias. The fitted parameters accurately predicted overall signal detection criterion, and its pupil response predicted shift (Fig. 3F,J; in humans, for all but the highest signal strength, which exhibited a

ceiling effect). Specifically, in both species, the starting point was strongly biased towards no-go irrespective of pupil response (Fig. 3C,G). Thus, overcoming this conservative choice bias required increasing their drift criterion. This increase in drift criterion occurred on trials with large pupil responses (Fig. 3D,E,H,I). We found either no, or a less consistent, relationship between pupil responses and drift rate, boundary separation, or non-decision time (Fig. S4).

Two observations verified that the drift criterion model used here well accounted for the overt behavior in the go/no-go task. First, as expected, drift rate increased with signal strength, reflecting the subjects' ability to accumulate strong sensory evidence more efficiently (Fig. S4A,G). Second, the fitted parameters accurately predicted overall RT and sensitivity, and, in mice, its pupil response predicted shift (Fig. S4).

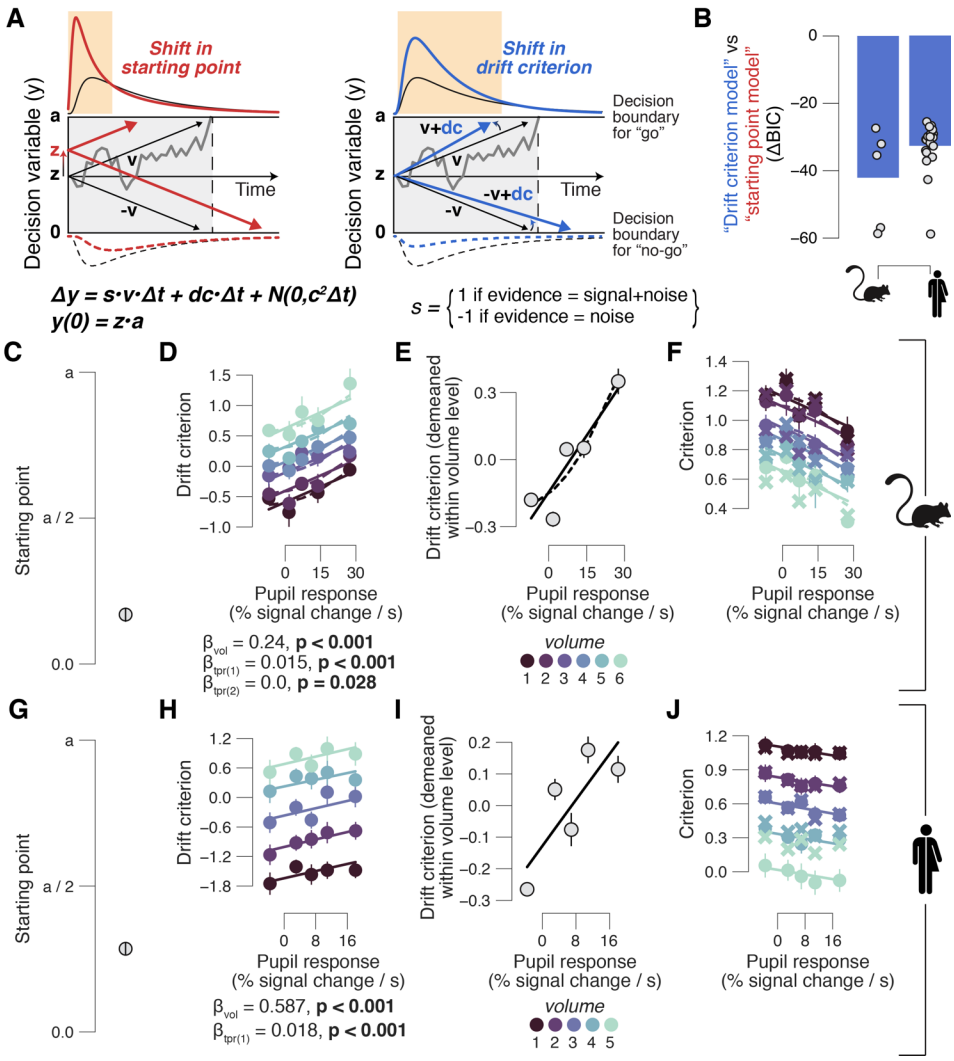

Figure 3. Phasic arousal selectively captures variation in the evidence accumulation process. (A) Schematic of drift diffusion model accounting for choices, and their associated RTs (for go-trials). Orange windows, RTs for which biased choices are expected under shifts in either "starting point" (*z*; left) or "drift criterion" (*dc*; right). Solid (dashed) lines, (implicit) RT distributions. **(B)** Bayesian information criterion (BIC) values for the drift criterion model versus the starting point model (Materials and Methods). Lower BIC values indicate a model that is better able to explain the data, after penalizing for complexity. Data points, individual subjects. **(C)** Starting point estimates in mice expressed as a fraction of the boundary separation (*a*). **(D)** Relationship between drift

criterion and task-evoked pupil response in mice, separately for the different signal strengths. Linear fits are plotted wherever the first-order fit was superior to the constant fit (Materials and Methods). Quadratic fits were plotted (dashed lines) wherever the second-order fit was superior to first-order fit. Stats, mixed linear modeling. **(E)** As D, but after demeaning drift criterion per signal strength. **(F)** Empirical (dots; as in Fig. 2A) and model-predicted (x-markers; Materials and Methods) choice bias as a function of pupil response in mice, separately for the different signal strengths. **(G-J)** As C-G, but for humans. C-J: group average (N = 5; N = 20); error bars, s.e.m.

The previous results suggest that the variations in systematic accumulation bias due to arousal would appear as random decision noise (i.e., drift rate variability) without taking trial-by-trial arousal responses into account. Indeed, this is what we found. We simulated RT distributions from two conditions that differed according to the fitted drift criterion (accumulation bias) estimates in the lowest and highest pupil-defined bin of each individual. As predicted, when fitting the model, drift rate variability was accurately recovered when drift criterion could vary with condition, but was significantly overestimated when drift criterion was fixed (Fig. 4).

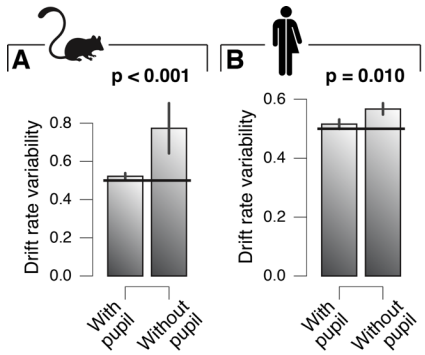

Figure 4. Variations in systematic accumulation bias due to arousal appears as decision noise. (A) Recovered drift rate variability for models with and without pupil predicted shift in drift criterion (Materials and Methods). Black line, true drift rate variability. **(B)** As panel A, but for humans. All panels: group average (N = 5; N = 20); error bars, s.e.m.; stats, paired-samples t-test.

Phasic arousal predicts a genuine bias suppression, not motor preparation

The central input to the pupil apparatus contains a transient at the subject's overt choice, as well as a sustained component throughout the decision interval (de Gee et al., 2017; 2014; Hupé et al., 2009; Murphy et al., 2016). In the go/no-task, these components (transient at lick / button-press and sustained preparatory activity) could have contributed to the pupil response amplitudes on go-trials but not (or less so) on no-go-trials, which in turn could trivially explain the relationship between pupil responses and choice bias. We eliminated contamination by the transient motor-related component by focusing on the initial pupil dilation in a protracted decision (Materials and Methods). However, it is possible that this did not fully account for the asymmetry between go- and no-go trials in terms of preparatory motor activity.

We thus asked humans (N = 24, 18 participants from go no-go experiment) to perform an auditory yes/no (forced-choice) detection task based on identical sensory stimuli to the go/no-go task. In this task, motor responses (and associated preparatory activity) were balanced across yes- and no-choices (Fig. 5A). Consistent with our go/no-go results in mice and humans, we observed that the pupil response predicted suppression of maladaptive perceptual choice bias (Fig. 5B). Furthermore, this bias suppression was, again, best explained by a change in drift criterion instead of starting point (Fig. 5C-E) and pushed behavior to a more optimal regime (Fig. S5A). Importantly, pupil response amplitudes

in the go/no-go and yes/no tasks were correlated across eighteen human subjects who participated in both experiments (Fig. 5F). This was true for yes- as well as no-choices. Therefore, the suppression of choice bias in our results does not reflect motor preparation.

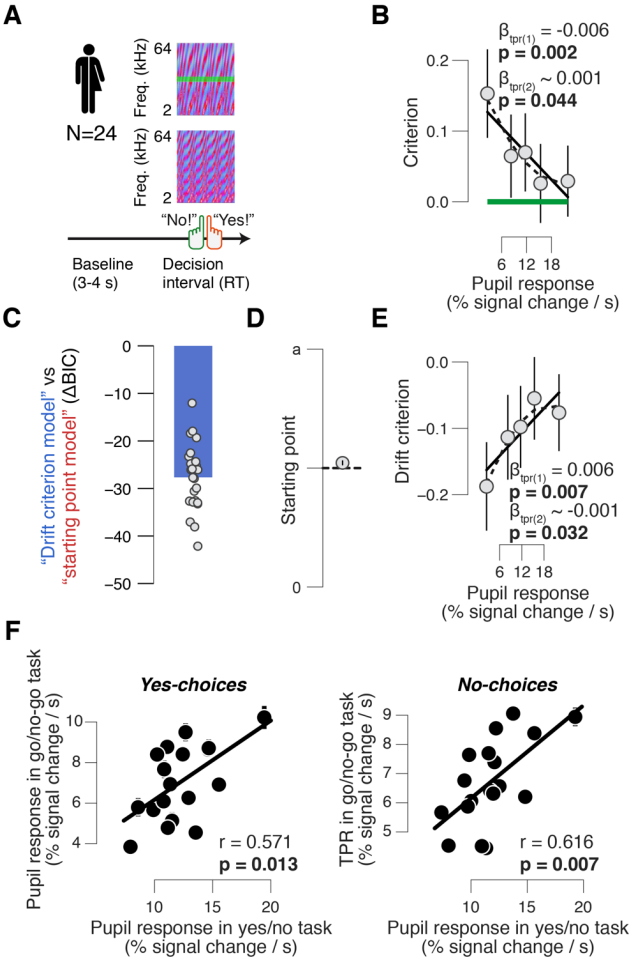

Figure 5. Phasic arousal predicts reduction of perceptual bias during yes/no decisions. (A) Auditory yes/no (forced choice) tone-in-noise detection task. Schematic sequence of events during a trial. Subjects reported the presence or absence of a faint signal (pure sine wave) embedded in noise (Materials and Methods). **(B)** Relationship between choice bias (quantified by signal detection criterion) and task-evoked pupil response. Linear fits were plotted if first-order fit was superior to constant fit; quadratic fits (dashed lines) where plotted if second-order fit was superior to first-order fit (Materials and Methods). Stats, mixed linear modeling. Green line, optimal criterion (Fig S5A). **(C)** Bayesian information criterion (BIC) values for the drift criterion model versus the starting point model (Materials and Methods). Lower BIC values indicate a model that is better able to explain the data, after penalizing for complexity. Data points, individual subjects. **(D)** Starting point estimates expressed as a fraction of the boundary separation *a*. **(E)** As B, but for drift criterion. **(F)** Left: individual task-evoked pupil response amplitude for yes-choices in the go/no-go task, plotted against individual pupil response amplitude for yes-choices in the yes/no (forced choice) task. Data points, individual subjects. Right: as left, but for no-choices. Stats, Pearson's correlation. B-E: group average (N = 24); error bars, s.e.m.

Arousal-linked bias reduction generalizes to high-level decisions

Does the arousal-related suppression of accumulation bias generalize to higher-level tasks, such as those in which psychology and behavioral economics have identified systematic deviations of human decision-making from rational choice (Kahneman and Tversky, 1979; Tsetsos et al., 2012; 2016; Tversky and Kahneman, 1974)? We studied one form of such bias: risk-seeking. To this end, we used a well-controlled task that emulates real-life decisions based on time-varying value information, such as deciding which of two fluctuating stock options had the higher returns in the past year (Tsetsos et al., 2016; 2012). Participants (N = 37) were instructed to average two sequences of payoff values (5–8 pairs) and, after the appearance of a response cue, choose the most "profitable" sequence (Fig. 6A; Materials and Methods). Because the decision is based on the accumulation of fluctuating samples (numbers), it lends itself naturally to the sequential sampling modeling approach we applied to perceptual decision-making before. Going beyond standard perceptual tasks, we specifically designed two trial types to quantify subjects' attitudes towards risk. On "narrow-correct" trials, the standard

deviation of the more profitable sequence was lower than that of the losing sequence; on "narrow-error", trials this was reversed (Fig. 6B and see Materials and Methods). Attitude towards risk was operationalized as "pro-variance bias": the fraction of high-variance choices pooled across both trial types (Materials and Methods). As in previous work (Tsetsos et al., 2016; 2012), subjects exhibited a systematic "pro-variance" bias indicating risk seeking (fraction of high-variance choices larger than 0.5; Fig. 6C).

Indeed, the pro-variance bias was suppressed as a function of pupil response (Fig. 6C). This bias was most reduced on intervals characterized by relatively strong pupil-linked phasic arousal responses. Pupil responses did not predict changes in RT or accuracy (Fig. S6C). We again used sequential sampling modeling to pinpoint the source of the pupil-linked pro-variance bias suppression. We fitted a previously introduced model that accounts for several key features of behavior in the current task (Tsetsos et al., 2016). The model entails the accumulation of the presented pairs of numbers across the trial. The accumulation is biased by a so-called selective gain parameter that prioritizes the (i.e. assigns larger weight to) the number that is higher on a given pair (Materials and Methods). This parameter gives rise to the pro-variance bias observed in choice behavior (Fig. S6D). Consistent with previous studies, selective gain in this dataset was larger than 0 (Fig. 6E), indicating an overall tendency towards down-weighting samples that were momentarily lower in value. But, critically, selective gain was pushed closer towards zero on trials characterized by large pupil responses (Fig. 6E), indicating a reduction in the pro-variance bias. We did not find robust evidence for pupil response predicting changes in other model parameters such as leak (controlling the time-constant of accumulation) or noise, although the latter exhibited a trend towards a systematic reduction as function of pupil response (Fig. SE). In other words, also in this high-level task did phasic pupil-linked arousal predict a selective change in evidence accumulation process, here reducing a risk-seeking bias.

In this task, a pro-variance bias can in fact be adaptive (i.e. improve reward rate) when noise corrupts the accumulation process downstream the representation of the incoming numbers: For a given level of accumulation noise the selective gain parameter maximizing accuracy is generally non-zero (Tsetsos et al., 2016). We used the best-fitting accumulation noise for each participant to obtain the selective gain parameter that maximized accuracy and calculated the pro-variance bias predicted by this selective gain parameter (green line in Fig. 6C for an across-participants average). The stronger the pupil responses, the closer to optimal the measured pro-variance bias, with the optimal pro-variance bias obtained for the largest pupil response bin. Therefore, consistent with the perceptual tasks, stronger phasic pupil arousal was associated with more optimal decision-making in a high-level task.

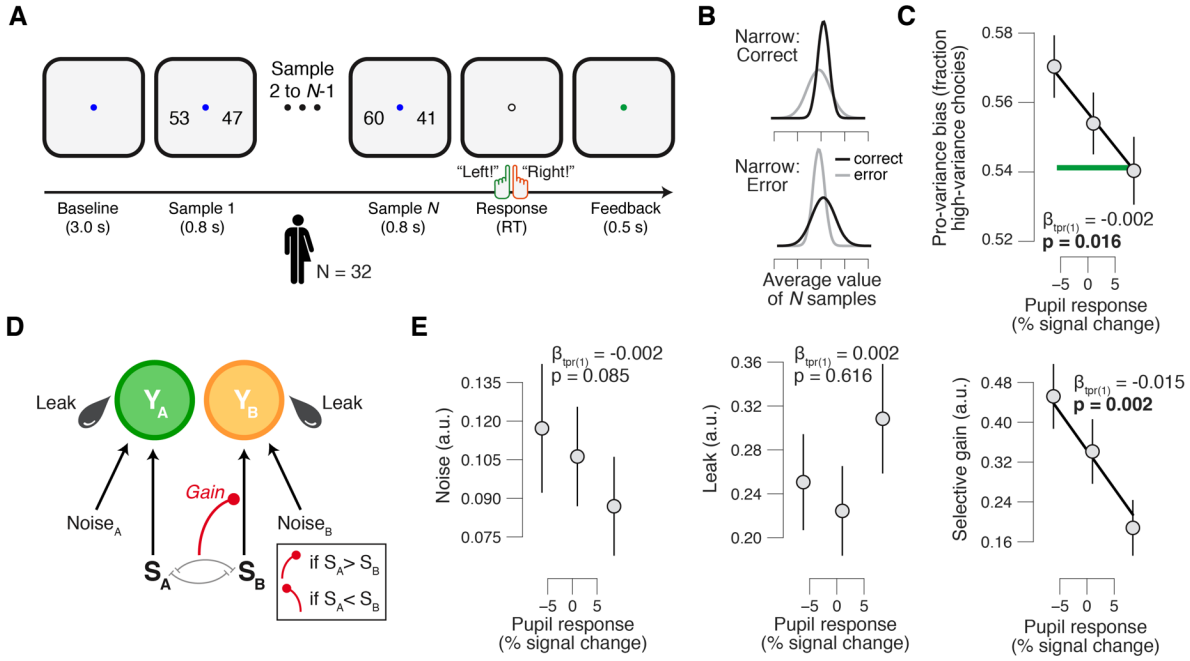

Figure 6. Phasic arousal predicts reduced risk seeking bias in a value-based choice task. (A) Value-based choice task. Schematic sequence of events during a trial. Subjects integrated two value sequences (5–8 pairs) and judged which sequence had on average the higher value. **(B)** On "narrow-correct" trials, the standard deviation of the more profitable sequence was lower than that of the losing sequence; on "narrow-error", trials this was reversed. **(C)** Relationship between pro-variance bias (Materials and Methods) and pupil response. Green line, optimal fraction of high-variance choices. Linear fits are plotted wherever the first-order fit was superior to the constant fit (Materials and Methods); quadratic fits were not superior to first-order fits. **(D)** Schematic of selective integration model (Materials and Methods). At the end of the accumulation period (after 5–8 number pairs) the accumulator with the higher total integrated value determines choice. The momentarily higher values pass onto the accumulation layer unaffected, but the momentarily lower ones are truncated according to the "selective gain" (*w*; Fig. S6E). **(E)** As panel C, but for the selective integration model parameters noise (left) and leak (middle) and selective gain (right). All panels: group average (N = 32); error bars, s.e.m.; stats, mixed linear modelling.

The phasic arousal-related bias suppression is distinct from ongoing arousal fluctuations

A concern might be that the bias suppression effects under large pupil responses reported here were due to associations between the preceding baseline pupil diameter and behavior. Such baseline effects might be "inherited" by the phasic pupil response through its commonly negative correlation with baseline pupil diameter (de Gee et al., 2014; Gilzenrat et al., 2010). This scenario could not account for the results reported here, for a number of reasons. First, in the go/no-go data sets, pupil responses were quantified as the rising slopes (see above), and those exhibited a negligible correlation to the preceding baseline diameter (mice: $r = 0.014$, $p = 0.742$; humans: $r = -0.041$, $p = 0.037$). Second, there was a non-monotonic association between baseline pupil diameter and decision bias in mice (McGinley et al., 2015a), in contrast to the monotonic pattern we report here for phasic arousal (Fig. 2B). Third, while the pupil responses were negatively related to baseline pupil diameter in the yes/no ($r = -0.159$, $p < 0.001$) and numbers data sets ($r = -0.482$, $p < 0.001$), in none of the human data sets was there any systematic (linear or non-monotonic) association between baseline pupil diameter and decision bias (go/no-go: $p = 0.975$; yes/no: $p = 0.557$; numbers: $p = 0.289$). Thus, the behavioral correlates of pupil

responses reported in this paper reflect genuine effects of phasic arousal uncontaminated by the baseline arousal state.

DISCUSSION

Owing to its widespread projections, the LC might profoundly impact cortical computation and behavior. The functional role of the task-evoked LC responses, however, has remained debated. We here established a new principle of this function, generalizing across species (humans and mice) and behavioral tasks (from perceptual to high-level decisions): task-evoked phasic arousal suppresses biases in evidence accumulation. We identified this principle by combining pupillometry and computational modeling of human and mouse choice behavior in the same auditory decision task. Task-evoked pupil responses occurred early during decision formation, even on trials without any motor response, and predicted a suppression of conservative choice bias. Behavioral modeling revealed that the bias reduction was due to a selective interaction with the evidence accumulation process that mediates the perceptual interpretation of the fluctuating sensory input. This result was consolidated in a forced-choice version of the auditory decision task, ruling out motor confounds. Finally, we established the pupil-linked suppression of evidence accumulation bias also for a higher-level form of human bias widely known in behavioral economics: risk-seeking. We conclude that the ongoing deliberation culminating in a choice (Shadlen and Kiani, 2013) is shaped by neuromodulatory responses that cause transient boosts in the global arousal state of the brain.

We chose the drift diffusion model to model the overt behavioral data of the go/no-go and yes/no tasks because the model (i) is sufficiently low-dimensional so that its parameter estimates are well constrained by the shape of RT distributions and choices, (ii) it has been shown to successfully account for behavioral data from a wide range of tasks, including go/no-go tasks (Ratcliff et al., 2016; Ratcliff and McKoon, 2008), and (iii) it is, under certain parameter regimens, equivalent to a reduction of biophysically detailed neural circuit models of decision-making (Bogacz et al., 2006; Wong and Wang, 2006). This choice of model required us to make three main assumptions. First, in the go/no-go task, participants accumulated the auditory evidence *within* each discrete noise sound during a trial, resetting this accumulation process before each next discrete sound. Second, in both the go/no-go and yes/no tasks, that subjects actively accumulated evidence towards both yes- and no-choices, and idea that is supported by neurophysiological data from yes/no tasks (Deco et al., 2007; Donner et al., 2009). Third, in the go/no-go task, subjects set an implicit boundary for no-choices (Ratcliff et al., 2016). While some of these assumptions (especially for go/no-go) should be validated in future physiological experiments, the quality of our model fits to the current data strongly suggest that the model was successful in accounting for the measured behavior and lend support in our conclusions drawn from the parameter estimates.

One important insight of the current work is that, without monitoring those phasic arousal responses, dynamic variations in systematic accumulation biases will be mistaken as random trial-to-trial variability in formal analyses of the decision process. Intrinsic variability is a pervasive feature of decision-making under uncertainty (Glimcher, 2005; Gold and Shadlen, 2007; Shadlen et al., 1996; Sugrue et al., 2005; Wyart and Koechlin, 2016). Most current models of decision-making account for this intrinsic behavioral variability by “noise” terms, such as the drift rate variability in the diffusion model (Bogacz et al., 2006; Ratcliff and McKoon, 2008). Such intrinsic variability may, however, originate from systematic biases in the inference process (Beck et al., 2012; Wyart and Koechlin, 2016). Here, we show that one part of these hidden biases is in fact dynamic, varying from trial to trial

as a function of arousal state, and can be tracked through pupil dilation. We further show that this accounts for a substantial portion of the drift rate variability.

Our insights into the function of phasic arousal stand in striking contrast to the recent observation of non-monotonic (inverted U) effects of tonic arousal levels on cortical computation and behavior – specifically the signal-to-noise ratio of sensory cortical responses and perceptual sensitivity (Gelbard-Sagiv et al., 2018; McGinley et al., 2015a). Instead, the relationship between phasic arousal and choice behavior was predominantly monotonic (approximately linear) and largely confined to decision bias. Importantly, our current work enables a direct head-to-head comparison of these functional correlates of tonic and phasic arousal within the exact same data set (mice). A previous report on that data set showed that, at intermediate arousal (medium baseline pupil size), the mice's behavioral performance was most rapid, accurate, and the least biased (McGinley et al., 2015a). In contrast, we here show that their behavioral performance was linearly related to phasic arousal, with the most rapid, accurate and least biased choices for the largest phasic arousal transients. It is tempting to speculate that these differences reflect different neuromodulatory systems governing tonic and phasic arousal. Indeed, rapid dilations of the pupil (phasic arousal) are tightly associated with phasic activity in noradrenergic axons, whereas slow changes in pupil (tonic arousal) are accompanied by sustained activity in cholinergic axons (Reimer et al., 2016).

Our data supports the hypothesis that phasic arousal interacts with the decision system as it unfolds, and thus, the intra-decisional account of task-evoked neuromodulatory responses (Dayan and Yu, 2006), as opposed to the post-decisional account (Aston-Jones and Cohen, 2005). These two competing accounts have emerged from empirical work conducted in different species and with different task designs, complicating direct comparisons. Our study involved the same task, physiological measurements, and model-based analysis approaches in mice and humans. This revealed remarkably consistent cross-species behavioral correlates of task-evoked arousal. Thus, we infer that the above discrepancy is not due to genuine differences in the behavioral impact of phasic arousal in different species.

Task-evoked neuromodulatory responses and the decision process in cortical circuits likely interact in a recurrent fashion. One possibility is that neuromodulatory responses alter the balance between "bottom-up" and "top-down" signaling across the cortical hierarchy (Friston, 2010; Gil et al., 1997; Hsieh et al., 2000; Kimura et al., 1999; Kobayashi et al., 2000). Here, bottom-up likelihood signals are sent by sensory cortical regions to downstream regions (e.g., encoding target presence in the detection tasks), while top-down signals encoding participants' prior beliefs about the state of the environment represented in these higher cortical regions are sent back to the lower levels of the hierarchy (Friston, 2010; Pouget et al., 2013). In our detection tasks, the prior may have been the absence of the target. Neuromodulators might dynamically reduce the weight of this prior in the inference process (Friston, 2010; Moran et al., 2013), in turn reducing the associated conservative bias in choice behavior. Critically, the amplitude of neuromodulator release might scale with the uncertainty (complement to confidence, or precision) about the incoming sensory data (Friston, 2010; Moran et al., 2013). Such an "inference-informed" neuromodulator release could be implemented by top-down control of the cortical systems for inference and decision-making over neuromodulatory brainstem centers, in line with anatomical connectivity (Aston-Jones & Cohen, 2005; Sara, 2009).

A related scenario that has been specifically proposed for phasic LC responses is that cortical regions driving LC-responses (e.g., the ACC) continuously compute the ratio of the posterior probability of the state of the world (target presence in the detection tasks) over its (estimated) prior probability (Dayan and Yu, 2006). When the prior is a conservative bias for "no", the LC is activated

when neural activity ramps towards the non-default "yes"-bound. The resulting LC activity might reset its cortical target circuits (Bouret and Sara, 2005) and override the default state (Dayan and Yu, 2006), facilitating the transition of the cortical decision circuitry towards the non-default state.

The finding that phasic arousal also optimizes choice behavior in the value-based choice task complements recent insights into the impact of tonic arousal and stress on value-based decision-making (Porcelli and Delgado, 2017). For example, acute stress reduces risk-seeking when making decisions involving financial gains (Porcelli and Delgado, 2009), it increases overexploitation in sequential foraging decisions (Lenow et al., 2017), and it impairs "model-based" goal-directed choice behavior (Otto et al., 2013; Schwabe et al., 2011). A direct comparison between these studies and ours is complicated by the different tasks used as well as the different behavioral states assessed (stress vs. phasic arousal). But the effects of acute stress on cognition and decision-making are mediated, at least in part, by tonic noradrenaline and dopamine release (Arnsten, 2015). It is tempting to interpret the current findings as a flip-side of the impairment in choice optimality found in the previous stress work: catecholaminergic modulation not only hampers, but can also boost, choice optimality when its duration is more confined.

Our study showcases the value of comparative experiments in humans and non-human species. It is widely appreciated that neuroscience can benefit from integrative approaches that bridge between different species and levels of analysis of brain function and behavior (Badre et al., 2015; Carandini, 2012; Krakauer et al., 2017; Sweis et al., 2018). Such approaches are challenging, however, hampered by possible species differences in behavioral strategies and the underlying neural computations, as well as by the incomplete understanding of the physiological relationship between brain signals obtained at different spatial scales (e.g., single neurons vs. neurophysiological signals vs. BOLD-fMRI responses (Donner and Siegel, 2011)). We propose that this challenge can be met by using physiological "reference signals" that serve as vehicles for linking of neural data across different scales and species. The identical behavioral correlates of task-evoked pupil dilation we report here for mice and humans, imply that pupil dilation can be used as a reference signal, at least in the decision-making tasks studied here. Future work should build on this finding in order to relate arousal-dependent variations in neural signatures of evidence accumulation obtained at different scales in humans, monkeys, and rodents (Brody and Hanks, 2016; Donner et al., 2009; Gold and Shadlen, 2007; Siegel et al., 2011).

In conclusion, our findings point to an intricate interplay between changes in the global arousal state of the brain and decision making in the face of uncertainty. Our results indicate that pupil-linked phasic arousal suppresses choice biases in evidence accumulation across species and across domains of decision-making.

MATERIALS AND METHODS

Subjects

We analyzed two previously unpublished human data sets, and re-analysed a previously published mice data set (McGinley et al., 2015a) and human data set (de Gee et al., 2017). All procedures concerning the animal experiments were carried out in accordance with Yale University Institutional Animal Care and Use Committee, and are described in detail elsewhere (McGinley et al., 2015a). Human subjects were recruited and participated in the experiment in accordance with the ethics committee of the Department of Psychology at the University of Amsterdam (go/no-go and yes/no tasks) or the ethics committee of the University Medical Center Hamburg-Eppendorf (value-based choice task). Human subjects gave written informed consent and received credit points (go/no-go and

yes/no tasks) or a performance-dependent monetary remuneration (value-based choice task) for their participation.

Five mice (all males; age range, 2–4 months) and twenty human subjects (15 females; age range, 19–28 y) performed the go/no-go task. Twenty-four human subjects (of which 18 had already participated in the go/no-go task; 20 females; age range, 19–28 y) performed an additional yes/no task. Thirty-seven human subjects (18 females; age range, 20–36 y) performed a value-based choice task, of which five were excluded from the analyses due to bad eye data quality and/or excessive blinking.

For the go/no-go task, mice performed between five and seven sessions (described in (McGinley et al., 2015a)), yielding a total of 2469–3479 trials per subject. For the go/no-go task, human participants performed 11 blocks of 60 trials each (distributed over two measurement sessions), yielding a total of 660 trials per participant. For the yes/no task, human participants performed between 11 and 13 blocks of 120 trials each (distributed over two measurement sessions), yielding a total of 1320–1560 trials per participant. For the value-based choice task, we only analyzed data from one of the experimental conditions ("pro-variance" condition, recorded during the placebo and nocebo sessions, see below) yielding a total of 288 trials per participant.

Behavioral tasks

Auditory go/no-go tone-in-noise detection task

Each trial consisted of two to seven consecutive distinct auditory noise stimuli (stimulus duration, 1 s; inter-stimulus-interval, 0.5 s). A weak signal tone (pure sine wave) was superimposed onto the last noise stimulus (Fig. 1A). The number of noise stimuli, and thus the signal position in the sequence, was randomly drawn beforehand. The probability of a target signal decreased linearly with sound position in the sequence (Fig. 1B), so as to keep hazard rate of signal onset approximately flat across the trial. Each trial was terminated by the subject's go-response (hit or false alarm) or after a no-go error (miss). Each interval consisted of only an auditory noise stimulus (McGinley et al., 2015a), or a pure sine wave (2 KHz) superimposed onto the noise. In the mice experiment, auditory stimuli were presented at an intensity of ~55dB. In the human experiment, auditory stimuli were presented at an intensity of 65dB using an IMG Stageline MD-5000DR over-ear headphone, suppressing ambient noise. Otherwise, it was the same set-up as in (de Gee et al., 2014).

Mice learned to respond during the signal+noise intervals and to withhold responses during noise intervals through training. Human participants were instructed to do the same. Mice responded by licking for sugar water reward. Humans responded by pressing a button with their right index finger. Correct yes-choices (hits) were followed by positive feedback: 4 uL of sugar water in the mice experiment, and a green fixation dot in the human experiment. False alarms were followed by an 8 s timeout. Humans received an 8 s timeout after misses too.

Target signal volume was randomly selected on each trial, under the constraint that each of six (mice) or five (humans) levels would occur equally often within each session (mice) or block of 60 trials (humans). The corresponding signal strengths exhibited a robust effect on mean accuracy, with highest accuracy for the loudest signal volume: $F(5,20) = 23.95$, $p < 0.001$) and $F(4,76) = 340.9$, $p < 0.001$), for mouse and human subjects respectively.

Auditory yes/no (forced choice) tone-in-noise detection task

Each trial consisted of two consecutive intervals (Fig. 5A): (i) the baseline interval (3-4 s uniformly distributed); (ii) the decision interval, the start of which was signaled by the onset of the auditory

stimulus and which was terminated by the subject's response (or after a maximum duration of 2.5 s). The decision interval consisted of only an auditory noise stimulus (McGinley et al., 2015a), or, on 50% of trials, a pure sine wave (2 KHz) superimposed onto the noise. Auditory stimuli were presented at the same intensity of 65dB using the same over-ear headphone as in the go/no-go task.

Participants were instructed to report the presence or absence of the signal by pressing one of two response buttons with their left or right index finger, once they felt sufficiently certain (free response paradigm). The mapping between perceptual choice and button press (e.g., "yes" –> right key; "no" –> left key) was counterbalanced across participants. After every 40 trials subjects were informed about their performance.

Throughout the experiment, the target signal volume was fixed at a level that yielded about 75% correct choices. Each participant's individual volume was determined before the main experiment using an adaptive staircase procedure (Quest). For this, we used a two-interval forced choice variant of the tone-in-noise detection yes/no task (one interval, signal+noise; the other, noise), in order to minimize contamination of the staircase by individual bias (generally smaller in two-interval forced choice than yes/no tasks). Indeed, the resulting threshold volumes produced a mean accuracy of 74.14% correct (±0.75 % s.e.m.) in the main experiment, very close to the targeted 75% correct level.

Value-based choice task

Each trial consisted of four consecutive intervals (Fig. 4A): (i) a pre-stimulus baseline interval (3.0 s); (ii) a stimulus interval consisting of 5–8 pairs of numbers; (iii) a response interval which was prompted by the fixation dot turning white and which was terminated by the participant's response or after a maximum of 2 s. Immediately after the response the fixation dot turned green or red, for correct and incorrect responses respectively, and stayed on screen for an additional 0.5 s (iv).

Participants were instructed to report which sequence (left or right) had, on average, the higher value. They indicated this judgment by pressing one of two response buttons, with the index finger of the left or right hand. Subjects received feedback at the end of each trial (green fixation dot, correct; red fixation dot, error). Participants were informed about their accuracy so far at the end of each block. On each session, participants received a maximum of €10 bonus (calculated as (X-0.7) x 10, where X was their overall fraction of correct choices or accuracy; for X > 0.8 the bonus was capped at €10).

The 5–8 pairs of 2-digit numerical values were black and presented sequentially, to the left and right of a central fixation point (0.34º diameter) against a grey background. Each number pair faded-in, changing linearly from grey to black for the first 300 ms, remained black for 200 ms, and then faded-out to grey for the last 300 ms. The viewing distance was 65 cm and each numerical character was 0.66º wide and 0.95º long.

In all trials there was a correct answer, with the average difference between the higher and the lower sequence being sampled from $d \sim U(1,12)$ with a mean of 6.5. This experiment contained three conditions, which were intermixed within a block of trials: a neutral condition, a condition designed to induce a "pro-variance" effect, and a condition designed to induce a "frequent winner" effect. In this report, we present analyses of the pro-variance condition; results of the neutral and frequent winner conditions will be the focus of another report. The pro-variance condition involved two types of trials, "narrow-correct" trials and "narrow-error" trials. In both types of trials the sequences were generated from Gaussian distributions, with the mean of the higher sequence (μH) sampled from $\mu H \sim U(45,65)$. The mean of the lower sequence was $\mu L = muh - d$. In the narrow-correct trials, the standard deviation of the higher sequence was $\sigma H = 10$ while the standard deviation of the lower sequence was $\sigma L = 20$; in the narrow-error trials this was reversed ($\sigma H = 20$ and $\sigma L = 10$).

This experiment was part of a larger study that also included MEG measurements of cortical activity combined with pharmacological intervention. Subjects performed the number intregration task in three measurement sessions (nocebo, placebo, drug [*lorazepam*]); they received an additional fixed €25 in the nocebo session, and an additional €70 in the placebo and drug sessions.

Eye data acquisition

The mice eye data acquisition is described elsewhere (McGinley et al., 2015a). The human experiments were conducted in a psychophysics laboratory (go/no-go and yes/no tasks) or in the MEG laboratory (value-based choice task). The left eye's pupil was tracked at 1000 Hz with an average spatial resolution of 15 to 30 min arc, using an EyeLink 1000 Long Range Mount (SR Research, Osgoode, Ontario, Canada), and it was calibrated once at the start of each block.

Analysis of task-evoked pupil responses

Preprocessing

Periods of blinks and saccades were detected using the manufacturer's standard algorithms with default settings. The remaining data analyses were performed using custom-made Python scripts. We applied to each pupil timeseries (i) linear interpolation of missing data due to blinks or other reasons (interpolation time window, from 150 ms before until 150 ms after missing data), (ii) low-pass filtering (third-order Butterworth, cut-off: 6 Hz), (iii) for human pupil data, removal of pupil responses to blinks and to saccades, by first estimating these responses by means of deconvolution and then removing them from the pupil time series by means of multiple linear regression (Knapen et al., 2016), and (iv) conversion to units of modulation (percent signal change) around the mean of the pupil time series from each measurement session. We computed the first derivative of the pupil size time series as:

Eq.1:

$$p'_n = p_{n+1} - p_n$$

with p as the preprocessed pupil size time series, and n as the index of each sample. Finally, the resulting pupil derivate time series were smoothened with a sliding boxcar car window (width, 50 ms).

Quantification of task-evoked pupil responses

The yes/no task was analogous in structure to the tasks from our previous work pupillometric work on decision-making (de Gee et al, 2014; 2017). We here computed task-evoked pupil responses as the maximum of the pupil derivative time series (Reimer et al., 2016) in the 500 ms before button press (grey windows in Fig. S5B). We used motor response-locking because motor responses, which occurred in all trials, elicit a transient pupil dilation response (de Gee et al., 2014; Hupé et al., 2009). Thus, locking pupil responses to the motor response balanced those motor components in the pupil responses across trials, eliminating them as a confounding factor for estimates of phasic arousal amplitudes. The resulting pupil bins were associated with strongly different overall pupil response amplitudes across the whole trial (Fig. S5D).

The go/no-go and value-based choice task entailed several deviations from the above task structure that posed different challenges for the quantification of task-evoked pupil responses. We met those by tailoring the analysis to the specifics of these task protocols, as described next. The go/no-task entailed, by design, an imbalance of motor responses between trials ending with different decisions, with no motor response for (implicit) no-choices Thus, the above-described transient motor component in the pupil response would yield larger pupil responses for yes- than for no-choices, even without any link between phasic arousal and decision bias. We took two measures to minimize contamination by this

motor imbalance. First, we quantified the pupil responses as the maximum of the pupil derivative in an early window that ranged from the start of the pupil derivative time course being significantly different from zero up to the first peak (grey windows in Fig. 1D,E). Second, we excluded decision intervals with a motor response before the end of this window plus a 50 ms buffer (Fig. S1E,F). In both species, the resulting pupil derivate defined bins were associated with strongly different overall pupil response amplitudes across the whole trial (Fig. S1B,D).

In the value-based choice task, the trials were substantially longer than in the go/no-go and yes/no tasks (4.0–6.4 s vs. ~1 s): the length of the value sequences varied systematically between trials (5–8 pairs), and the high-contrast numbers elicited an initial constriction of the pupil (Fig. S6A, initial dip below pre-stimulus baseline level during the first 1.5 s). In order to quantify the amplitude of phasic arousal across the full interval of evidence accumulation, we computed pupil responses as the mean pupil size from 1.5 s to 4.5 s after the onset of the first pair of samples (grey window in Fig. S6A), with the pre-trial baseline pupil size (mean pupil size in the 500 ms before the first pair of samples) subtracted out. As pupil diameter increased with each sample after the first (Fig. S6A), larger pupil responses were to be expected for 8-sample compared to 5-sample trials. Therefore, we computed pupil responses aligned to stimulus onset, while excluding (i) the initial dip during after the first pair of samples (likely due to the pupil light reflex) and (ii) motor and/or feedback-related components occurring post 4.5 s for the shortest trials (5 samples) (Fig. S6A, left). The resulting pupil response defined bins were associated with strongly different overall pupil response amplitudes across the whole trial (Fig. S5D).

For analyses of the go/no-go and yes/no tasks, we used five equally populated bins of task-evoked pupil response amplitudes; we used three bins for the value-based choice task, in which subjects completed fewer trials.

Analysis and modeling of choice behavior

In the go/no-go task all stimuli in one trial as defined in the experiment (i.e., sequence discrete signal+noise or noise sounds) were analyzed as a separate decision. The first stimulus of each trial (see *Behavioral tasks*) was excluded from the analyses, because this interval served as a reference and never included the target signal (pure sine wave). In the go/no-go and yes/no tasks, reaction time (RT) was defined as the time from stimulus onset until the lick or button press. In the value-based choice task, RT was defined as the time from the last sample offset until the button press. In the mice go/no-go data set, intervals with RTs shorter than 240 ms were excluded from the analyses (see *Quantification of task-evoked pupillary responses* and Fig. S1E); in the human go/no-go data set, intervals with RTs shorter than 510 ms were excluded from the analyses (Fig. S1F).

Signal-detection theoretic modeling (go/no-go and yes/no tasks)

Signal detection metrics d' and criterion (Green and Swets, 1966) were computed separately for five pupil response defined bins. We estimated d' as the difference between z-scores of hit- and false-alarm rates. We estimated criterion by averaging the z-scores of hit- and false-alarm rates and multiplying the result by -1. In the go/no-go task, the same false alarm rate was used for each signal strength (difficulty level).

Determining optimal choice bias in the go/no-go task

We simulated 50000 trials for each combination of a range of perceptual sensitivities and choice biases (signal detection d' and criterion). Sensitivity ranged from 0.5 to 3.0 in steps of 0.5 and criterion ranged from -3.5 to 3.5 in steps of 0.05. On each trial, target signal position (#2-7 in the sequence) was determined as in the actual task (see above). On every sound interval, the agent's internal decision

variable (DV) was randomly drawn from a noise or signal+noise distribution which were d' apart. The noise distribution was three times larger than the signal+noise distribution because subjects encounter more noise sounds (follows from the probabilities in Fig. 1B). Every encountered noise sound added 1.5 s (1 s sound + 0.5 s ISI; see Fig. 1A) to total time. A correct reject (DV drawn from noise distribution < criterion) was followed by the next sound in the same sequence. A hit (DV drawn from signal+noise distribution > criterion) resulted in a reward and the completion of the trial. A false alarm (DV drawn from noise distribution > criterion) resulted in a timeout (additional 8 s added to total time) and the abortion of the trial without obtaining a reward. A miss (DV drawn from signal+noise distribution < criterion) resulted the abortion of the trial without obtaining a reward. For the human version of the go/no-go task, an additional 8 s was added to total time after misses. Optimality was defined as the criterion value that maximized reward rate (# rewards / total time).

Drift diffusion modeling (go/no-go and yes/no tasks)

We used the HDDM 0.6.1 package (Wiecki et al., 2013) to fit the drift diffusion model for each subject, each of five pupil response-bins and (in the go/no-go task), separately for each signal strength. We fitted the model using based on RT quantiles, using the so-called G square method (code pull-requested into the master HDDM repository on Github). The RT distributions for yes- and no-choices were represented by the 0.1, 0.3, 0.5, 0.7 and 0.9 quantiles, and, along with the associated response proportions, contributed to G square. In the go/no-go task, a single bin containing the number of no-go-choices contributed to G square (Ratcliff et al., 2016). Fitting the model to RT distributions for the separate responses (termed "stimulus coding" in (Wiecki et al., 2013)) enabled estimating parameters that could have induced biases towards specific choices.

We fitted two separate version of the drift diffusion model. A "starting point model" allowed the following parameters to vary with five pupil response-bins: (i) the starting point of the evidence accumulation process; (ii) the mean drift rate across trials; (iii) the separation between both decision bounds (i.e., response caution); (iv) the non-decision time (sum of the latencies for sensory encoding and motor execution of the choice). Additionally, drift criterion (an evidence independent constant added to the drift) and drift rate variability were fitted but fixed across the pupil-defined bins. The alternative "drift criterion model" was as the starting point model, except that starting point was fixed across the pupil-defined bins, and drift criterion was instead allowed to vary with pupil response-bin.

In the go/no-go task, to formally compare the starting point and drift criterion model based on Bayesian information criterion (BIC, see below), we first pooled data across signal strength. This was done to ensure the same number of parameters in starting point and drift criterion models (separate fits per signal strength would have required drift criterion to additionally vary with signal strength). Having established the drift criterion model as superior (Fig. 3B), we then refitted the drift criterion model, when additionally allowing drift rate and drift criterion to vary with signal strength, the same noise only intervals were re-used when fitting the model to each signal strength.

We used BIC to select the model which provided the best fit to the data (Schwarz, 1978). BIC compared models based on their maximized log-likelihood value, while penalizing for the number of parameters. Lower BIC values indicated a model that better explained the data, taking model complexity into account. BIC differences of 10 are generally taken as a threshold for considering one model a decisively better fit.

To verify that the drift criterion model indeed accounted for the pupil response-dependent changes in overt choice fractions (i.e., signal detection criterion, Fig. 3F,J), we simulated a new dataset using

the fitted drift diffusion model parameters. Separately per subject, we simulated 5000 trials for each signal strength and pupil bin, while ensuring that the fraction of signal+noise vs. noise trials matched that of the empirical data; we then computed signal detection criterion for every bin (as described above).

We used a similar approach to test if, without monitoring task-evoked pupil responses, systematic variations in accumulation bias (drift criterion) would appear as random trial-to-trial variability in the accumulation process (drift rate variability) (Fig. 4). For simplicity, we now pooled across signal strengths and simulated 50000 trials from two conditions that differed according to the fitted drift criterion (accumulation bias) estimates in the lowest and highest pupil-defined bin of each individual; drift rate, boundary separation and non-decision time were fixed to the mean across pupil bins of each individual; drift rate variability was fixed to 0.5. We then fitted the drift criterion model as described above to the simulated data, and another version of the model in which we fixed drift criterion across the two conditions.

Modeling behavior from the value-based choice task

The standard DDM model, in which the drift-rate is fixed across the trial, is not well-suited for the value-based choice task because it would ignore the fact that the decision input fluctuates widely (i.e. not due to internal noise). However, even when taking this non-stationarity into account, variants of the DDM model fall short of explaining the pro-variance bias. This point is extensively exposed in (Tsetsos et al., 2012).

Pro-variance bias (Tsetsos et al., 2016; 2012) was computed as the as the fraction of high-variance choices across both trial types (sequence A when $\sigma A > \sigma B$; or sequence B when $\sigma A < \sigma B$). We fitted a model dubbed "selective integration" (Tsetsos et al., 2016), which describes a decision-process based on the accumulation of two sequences of simultaneously presented inputs. The model assumes that the two sequences have the same number of inputs, and that each pair of inputs was presented at a discrete time-step for a fixed time interval. Additionally, based on the findings in (Tsetsos et al., 2016), we assumed that the inputs were not corrupted by noise prior to accumulation. The two sequences were labelled S_A and S_B, with $S_A(t)$ indicating the value of sequence A at the (discrete) sample t. Two accumulators (Y_A and Y_B) integrated the values of the sequences across time according to the following difference equations:

Eq.2:

$$\mathrm{Y}_A(t) = (1-\lambda)\cdot \mathrm{Y}_A(t-1) + \mathrm{I}_A(t) + \xi \cdot \xi_A(t)$$

$$\mathrm{Y}_B(t) = (1-\lambda)\cdot \mathrm{Y}_B(t-1) + \mathrm{I}_B(t) + \xi \cdot \xi_B(t)$$

with t as the current discrete time-step (or sample), λ as the leak in the accumulation, $I_{A,B}(t)$ as the inputs to the two accumulators on a given time-step, ξ as the standard deviation of the noise at the accumulation level, and $\zeta_{A,B}(t)$ as the standard Gaussian samples (independent from each other and across time). The accumulators were initialized at 0: $Y_A(0) = Y_B(0) = 0$. At the end of the accumulation period (at $t = T$, with T as the total number of pairs of samples presented) a decision was made in favor of the accumulator with the higher total integrated value. If both accumulators ended up with the same total integrated value, a decision was made randomly.

The inputs to the two accumulators, $I_{A,B}(t)$, reflected the modified sequence values after a selective integration filter is applied, referred to as "selective gain" and implemented as follows:

Eq.3:

$$\mathrm{I}_A(t) = \theta(\mathrm{S}_A(t), \mathrm{S}_B(t)) \cdot \mathrm{S}_A(t)$$

$$\mathrm{I}_B(t) = \theta(\mathrm{S}_B(t), \mathrm{S}_A(t)) \cdot \mathrm{S}_B(t)$$

with function θ as follows:

Eq.4:

$$\theta(\mathrm{x}, y) = \begin{cases} 1 \text{ if } x \geq y \\ 1 - w \text{ if } x < y \end{cases}$$

This function returned a value of 1 if the first argument was equal or larger than the second and a value *w* (selective gain parameter) otherwise.

We fitted a version of the model that had three free parameters: leak (λ), late noise (ξ) and selective gain (w). For a given parameter set and each experimental trial (i.e. using the exact numerical sequences that participants observed) we numerically obtained the choice probability of the selective integration model and the likelihood of the parameter set given the empirical choice on that trial (Tsetsos et al., 2016). We summed the negative log likelihoods across trials in order to obtain the overall negative log-likelihood. Minimizing the overall negative log-likelihood was done via a two-stage procedure: (i) by an initial grid search in the parameter space, and (ii) by feeding the 20 best fitting parameter sets obtained from the grid search, as starting points in a SIMPLEX optimization routine. The model was fitted to data from the pro-variance trials only and separately for each level of pupil response.

Determining optimal choice bias in the value-based choice task

To understand whether behavior approaches optimality with stronger pupil responses we fitted the data to all trials, regardless of pupil response. This allowed us to derive the per participant model parameters. Using the best-fitting noise and leak parameters we calculated the level of selective gain that achieves maximum accuracy. As shown in *(Tsetsos et al., 2016)*, the optimal level of selective gain increases with the level of late noise. Finally, we predicted optimal pro-variance bias (i.e., the bias maximizing percentage correct) using the optimal selective gain for each participant (Fig. 6C green line for an average).

Statistical comparisons

We used a mixed linear modeling approach implemented in the *R*-package *lme4* (Bates et al., 2015) to quantify the dependence of several metrics of overt behavior, or of estimated model parameters (see above), on pupil response. For the go/no-go task, we simultaneously quantified the dependence on signal strength. Our approach was analogous to sequential polynomial regression analysis (Draper and Smith, 1998), but now performed within a mixed linear modeling framework. In the first step, we fitted three mixed models to test whether pupil responses predominantly exhibited no effect (zero-order

polynomial), a monotonic effect (first-order), or a non-monotonic effect (second-order) on the behavioral metric of interest (y). The fixed effects were specified as:

Eq.5:

$$\text{Model 1: } y \sim \beta_0 1 + \beta_1 S$$

$$\text{Model 2: } y \sim \beta_0 1 + \beta_1 S + \beta_2 TPR^1$$

$$\text{Model 3: } y \sim \beta_0 1 + \beta_1 S + \beta_2 TPR^1 + \beta_3 TPR^2$$

with β as regression coefficients, S as the signal strength (for go/no-go task), and TPR as the bin-wise task-evoked pupil response amplitudes. We included the maximal random effects structure justified by the design (Barr et al., 2013). For data from the go/no-go task, the random effects were specified to accommodate signal strength coefficient to vary with participant, and the intercept and TPR-coefficients to vary with signal strength and participant. For data from the yes/no and value-based choice tasks, the random effects were specified to accommodate the intercept and TPR-coefficients to vary with participant. The mixed models were fitted through maximum likelihood estimation. Each model was then sequentially tested in a serial hierarchical analysis, based on chi-squared statistics. This analysis was performed for the complete sample at once, and it tested whether adding the next higher order model yielded a significantly better description of the response than the respective lower order model. We tested models from the zero-order (constant, no effect of pupil response) up to the second-order (quadratic, non-monotonic). In the second step, we refitted the winning model through restricted maximum likelihood estimation, and computed p-values with Satterthwaite's method implemented in the *R*-package *lmerTest* (Kuznetsova et al., 2017).

We used paired-sample t-tests to test for significant differences between the pupil derivative time course and 0, and between pupil response amplitudes for yes- versus no-choices.

Data and code sharing

The data are publicly available on [to be filled in upon publication]. Analysis scripts are publicly available on [to be filled in upon publication].

ACKNOWLEDGMENTS

We thank Daniëlle Rijkmans, Guusje Boomgaard and Christopher David Riddell for help with the data collection for the human auditory detection tasks, and all members of the Donner lab for discussion. This research was supported by the German Research Foundation (DFG, grant numbers: DO 1240/3–1 and SFB 936A7 to THD), European Commission CH2020 7th Framework Programme (Marie Skłodowska-Curie Individual Fellowship: 658581-CODIR, to KT and THD), and the National Institutes of Health (R03DC015618, to MJM).

AUTHOR CONTRIBUTIONS

JWdG, Conceptualization, Formal analysis, Investigation, Writing—original draft, Writing—review and editing; KT, Conceptualization, Formal analysis, Investigation, Writing—original draft, Writing—review and editing; DAM, Conceptualization, Writing—review and editing; MJM,

Conceptualization, Formal analysis, Investigation, Writing—original draft, Writing—review and editing; THD, Conceptualization, Writing—original draft, Writing—review and editing.

REFERENCES

Arnsten, A.F.T., 2015. Stress weakens prefrontal networks: molecular insults to higher cognition. Nat Neurosci 18, 1376–1385. doi:10.1038/nn.4087

Aston-Jones, G., Cohen, J.D., 2005. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance 28, 403–450. doi:10.1146/annurev.neuro.28.061604.135709

Badre, D., Frank, M.J., Moore, C.I., 2015. Interactionist Neuroscience. Neuron 88, 855–860. doi:10.1016/j.neuron.2015.10.021

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang 68, 255–278. doi:10.1016/j.jml.2012.11.001

Beck, J.M., Ma, W.J., Pitkow, X., Latham, P.E., Pouget, A., 2012. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron 74, 30–39. doi:10.1016/j.neuron.2012.03.016

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. Psychol Rev 113, 700–765. doi:10.1037/0033-295X.113.4.700

Bouret, S., Sara, S.J., 2005. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends Neurosci. 28, 574–582. doi:10.1016/j.tins.2005.09.002

Brody, C.D., Hanks, T.D., 2016. Neural underpinnings of the evidence accumulator. Curr. Opin. Neurobiol. 37, 149–157. doi:10.1016/j.conb.2016.01.003

Carandini, M., 2012. From circuits to behavior: a bridge too far? Nat Neurosci 15, 507–509. doi:10.1038/nn.3043

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., de Gardelle, V., Herce Castañón, S., Summerfield, C., 2014. Adaptive gain control during human perceptual choice. Neuron 81, 1429–1441. doi:10.1016/j.neuron.2014.01.020

Clayton, E.C., Rajkowski, J., Cohen, J.D., Aston-Jones, G., 2004. Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. J. Neurosci. 24, 9914–9920. doi:10.1523/JNEUROSCI.2446-04.2004

Dayan, P., Yu, A.J., 2006. Phasic norepinephrine: a neural interrupt signal for unexpected events. Network 17, 335–350. doi:10.1080/09548980601004024

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. doi:10.7554/eLife.23232

de Gee, J.W., Knapen, T., Donner, T.H., 2014. Decision-related pupil dilation reflects upcoming choice and individual bias. Proc. Natl. Acad. Sci. U.S.A. 111, E618–25. doi:10.1073/pnas.1317557111

Deco, G., Pérez-Sanagustín, M., de Lafuente, V., Romo, R., 2007. Perceptual detection as a dynamical bistability phenomenon: A neurocomputational correlate of sensation. PNAS 104, 20073–20077. doi:10.1073/pnas.0709794104

Donner, T.H., Siegel, M., 2011. A framework for local cortical oscillation patterns. Trends Cogn. Sci. (Regul. Ed.) 15, 191–199. doi:10.1016/j.tics.2011.03.007

Donner, T.H., Siegel, M., Fries, P., Engel, A.K., 2009. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. Curr. Biol. 19, 1581–1585. doi:10.1016/j.cub.2009.07.066

Draper, N.R., Smith, H., 1998. Applied Regression Analysis. Wiley-Interscience.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat Rev Neurosci 11, 127–138. doi:10.1038/nrn2787

Gelbard-Sagiv, H., Magidov, E., Sharon, H., Hendler, T., Nir, Y., 2018. Noradrenaline Modulates Visual Perception and Late Visually Evoked Activity. Curr. Biol. doi:10.1016/j.cub.2018.05.051

Gil, Z., Connors, B.W., Amitai, Y., 1997. Differential regulation of neocortical synapses by neuromodulators and activity. Neuron 19, 679–686. doi:10.1016/S0896-6273(00)80380-3

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. Cogn Affect Behav Neurosci 10, 252–269. doi:10.3758/CABN.10.2.252

Glimcher, P.W., 2005. Indeterminacy in brain and behavior. Annu Rev Psychol 56, 25–56. doi:10.1146/annurev.psych.55.090902.141429

Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu. Rev. Neurosci. 30, 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Green, D.M., Swets, J.A., 1966. Signal detection theory and psychophysics. 1966. New York. doi:10.1111/j.1467-985X.2005.346_1.x/abstract

Harris, K.D., Thiele, A., 2011. Cortical state and attention. Nat Rev Neurosci 12, 509–523. doi:10.1038/nrn3084

Hsieh, C.Y., Cruikshank, S.J., Metherate, R., 2000. Differential modulation of auditory thalamocortical and intracortical synaptic transmission by cholinergic agonist. Brain Res. 880, 51–64.

Hupé, J.-M., Lamirel, C., Lorenceau, J., 2009. Pupil dynamics during bistable motion perception. J Vis 9, 10. doi:10.1167/9.7.10

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, 221–234. doi:10.1016/j.neuron.2015.11.028

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47, 263. doi:10.2307/1914185

Kalwani, R.M., Joshi, S., Gold, J.I., 2014. Phasic Activation of Individual Neurons in the Locus Ceruleus/Subceruleus Complex of Monkeys Reflects Rewarded Decisions to Go But Not Stop. Journal of Neuroscience 34, 13656–13669. doi:10.1523/JNEUROSCI.2566-14.2014

Kimura, F., Fukuda, M., Tsumoto, T., 1999. Acetylcholine suppresses the spread of excitation in the visual cortex revealed by optical recording: possible differential effect depending on the source of input. Eur. J. Neurosci. 11, 3597–3609.

Knapen, T., de Gee, J.W., Brascamp, J., Nuiten, S., Hoppenbrouwers, S., Theeuwes, J., 2016. Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance. PLoS ONE 11, e0155574. doi:10.1371/journal.pone.0155574

Kobayashi, M., Imamura, K., Sugai, T., Onoda, N., Yamamoto, M., Komai, S., Watanabe, Y., 2000. Selective suppression of horizontal propagation in rat visual cortex by norepinephrine. Eur. J. Neurosci. 12, 264–272.

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., Poeppel, D., 2017. Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron 93, 480–490. doi:10.1016/j.neuron.2016.12.041

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest Package: Tests in Linear Mixed Effects Models. J. Stat. Soft. 82. doi:10.18637/jss.v082.i13

Lak, A., Nomoto, K., Keramati, M., Sakagami, M., Kepecs, A., 2017. Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. Curr. Biol. 27, 821–832. doi:10.1016/j.cub.2017.02.026

Larsen, R.S., Waters, J., 2018. Neuromodulatory Correlates of Pupil Dilation. Frontiers in Neural Circuits 12. doi:10.3389/fncir.2018.00021

Lee, C.R., Margolis, D.J., 2016. Pupil Dynamics Reflect Behavioral Choice and Learning in a Go/NoGo Tactile Decision-Making Task in Mice. Front Behav Neurosci 10, 200. doi:10.3389/fnbeh.2016.00200

Lee, S.-H., Dan, Y., 2012. Neuromodulation of brain states. Neuron 76, 209–222. doi:10.1016/j.neuron.2012.09.012

Lenow, J.K., Constantino, S.M., Daw, N.D., Phelps, E.A., 2017. Chronic and Acute Stress Promote Overexploitation in Serial Decision Making. J. Neurosci. 37, 5681–5689. doi:10.1523/JNEUROSCI.3618-16.2017

Liu, Y., Rodenkirch, C., Moskowitz, N., Schriver, B., Wang, Q., 2017. Dynamic Lateralization of Pupil Dilation Evoked by Locus Coeruleus Activation Results from Sympathetic, Not Parasympathetic, Contributions. Cell Rep 20, 3099–3112. doi:10.1016/j.celrep.2017.08.094

McGinley, M.J., David, S.V., McCormick, D.A., 2015a. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron 87, 179–192. doi:10.1016/j.neuron.2015.05.038

McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., Tolias, A.S., Cardin, J.A., McCormick, D.A., 2015b. Waking State: Rapid Variations Modulate Neural and Behavioral Responses. Neuron 87, 1143–1161. doi:10.1016/j.neuron.2015.09.012

Moran, R.J., Campo, P., Symmonds, M., Stephan, K.E., Dolan, R.J., Friston, K.J., 2013. Free energy, precision and learning: the role of cholinergic neuromodulation. J. Neurosci. 33, 8227–8236. doi:10.1523/JNEUROSCI.4255-12.2013

Murphy, P.R., Boonstra, E., Nieuwenhuis, S., 2016. Global gain modulation generates time-dependent urgency during perceptual choice in humans. Nat Commun 7, 13526. doi:10.1038/ncomms13526

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. Hum Brain Mapp 35, 4140–4154. doi:10.1002/hbm.22466

Otto, A.R., Raio, C.M., Chiang, A., Phelps, E.A., Daw, N.D., 2013. Working-memory capacity protects model-based learning from stress. PNAS 110, 20941–20946. doi:10.1073/pnas.1312011110

Parikh, V., Kozak, R., Martinez, V., Sarter, M., 2007. Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron 56, 141–154. doi:10.1016/j.neuron.2007.08.025

Porcelli, A.J., Delgado, M.R., 2017. Stress and Decision Making: Effects on Valuation, Learning, and Risk-taking. Current Opinion in Behavioral Sciences 14, 33–39. doi:10.1016/j.cobeha.2016.11.015

Porcelli, A.J., Delgado, M.R., 2009. Acute stress modulates risk taking in financial decision making. Psychol Sci 20, 278–283. doi:10.1111/j.1467-9280.2009.02288.x

Pouget, A., Beck, J.M., Ma, W.J., Latham, P.E., 2013. Probabilistic brains: knowns and unknowns. Nat Neurosci 16, 1170–1178. doi:10.1038/nn.3495

Rajkowski, J., Majczynski, H., Clayton, E., Aston-Jones, G., 2004. Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. J. Neurophysiol. 92, 361–371. doi:10.1152/jn.00673.2003

Ratcliff, R., Huang-Pollock, C., McKoon, G., 2016. Modeling Individual Differences in the Go/No-Go Task With a Diffusion Model. Decision. doi:10.1037/dec0000065

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput 20, 873–922. doi:10.1162/neco.2008.12-06-420

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. Nat Commun 7, 13289. doi:10.1038/ncomms13289

Schriver, B.J., Bagdasarov, S., Wang, Q., 2018. Pupil-linked arousal modulates behavior in rats performing a whisker deflection direction discrimination task. J. Neurophysiol. 23, 9565. doi:10.1152/jn.00290.2018

Schwabe, L., Höffken, O., Tegenthoff, M., Wolf, O.T., 2011. Preventing the stress-induced shift from goal-directed to habit action with a β-adrenergic antagonist. J. Neurosci. 31, 17317–17325. doi:10.1523/JNEUROSCI.3304-11.2011

Schwarz, G., 1978. Estimating the Dimension of a Model. The Annals of Statistics 6, 461–464. doi:10.1214/aos/1176344136

Shadlen, M.N., Britten, K.H., Newsome, W.T., Movshon, J.A., 1996. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. Journal of Neuroscience 16, 1486–1510.

Shadlen, M.N., Kiani, R., 2013. Decision making as a window on cognition. Neuron 80, 791–806. doi:10.1016/j.neuron.2013.10.047

Siegel, M., Engel, A.K., Donner, T.H., 2011. Cortical network dynamics of perceptual decision-making in the human brain. Front Hum Neurosci 5, 21. doi:10.3389/fnhum.2011.00021

Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2005. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci 6, 363–375. doi:10.1038/nrn1666

Sweis, B.M., Abram, S.V., Schmidt, B.J., Seeland, K.D., MacDonald, A.W., Thomas, M.J., Redish, A.D., 2018. Sensitivity to "sunk costs" in mice, rats, and humans. Science 361, 178–181. doi:10.1126/science.aar8644

Tsetsos, K., Chater, N., Usher, M., 2012. Salience driven value integration explains decision biases and preference reversal. Proc. Natl. Acad. Sci. U.S.A. 109, 9659–9664. doi:10.1073/pnas.1119569109

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., Summerfield, C., 2016. Economic irrationality is optimal during noisy decision making. Proc. Natl. Acad. Sci. U.S.A. 113, 3102–3107. doi:10.1073/pnas.1519157113

Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185, 1124–1131. doi:10.1126/science.185.4157.1124

Urai, A.E., de Gee, J.W., Donner, T.H., 2018. Choice history biases subsequent evidence accumulation. doi:10.1101/251595

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, 7866–7877. doi:10.1523/JNEUROSCI.0454-15.2015

Wiecki, T.V., Sofer, I., Frank, M.J., 2013. HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Front Neuroinform 7, 14. doi:10.3389/fninf.2013.00014

Wong, K.-F., Wang, X.-J., 2006. A recurrent network mechanism of time integration in perceptual decisions. J. Neurosci. 26, 1314–1328. doi:10.1523/JNEUROSCI.3733-05.2006

Wyart, V., Koechlin, E., 2016. Choice variability and suboptimality in uncertain environments. Current Opinion in Behavioral Sciences. doi:10.1016/j.cobeha.2016.07.003

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, 459–482. doi:10.1002/cne.920180503

SUPPLEMENTARY FIGURES

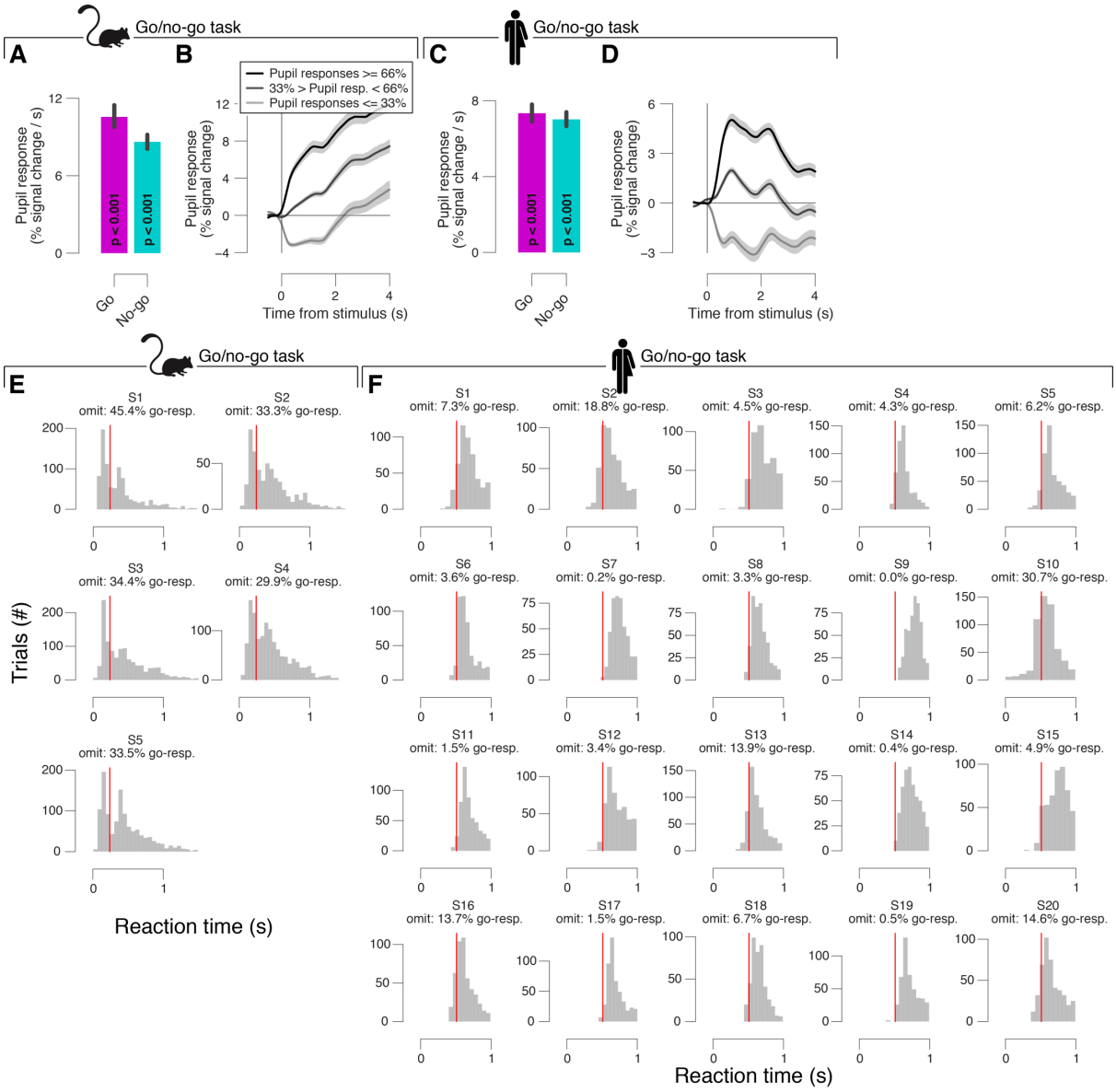

Figure S1. Quantifying pupil responses and behavior in mice and humans. (A) Task-evoked pupil responses in mice sorted into go- and no-go choices (pooled across signal strengths). Stats, paired-samples t-test. **(B)** Overall pupil response time courses in mice for three pupil derivative defined bins (pooled across signal strengths). **(C,D)** As A,B, but for humans. **(E)** RT distributions of individual mice. Red lines, group average latency of the first peak in pupil slope timeseries plus a 50 ms buffer, which was used as a cut-off for excluding decision intervals in order to control for a potential motor confound in our task-evoked pupil response measures (Materials and Methods). **(F)** As E, but for individual humans. All panels: error bars or shading, s.e.m.

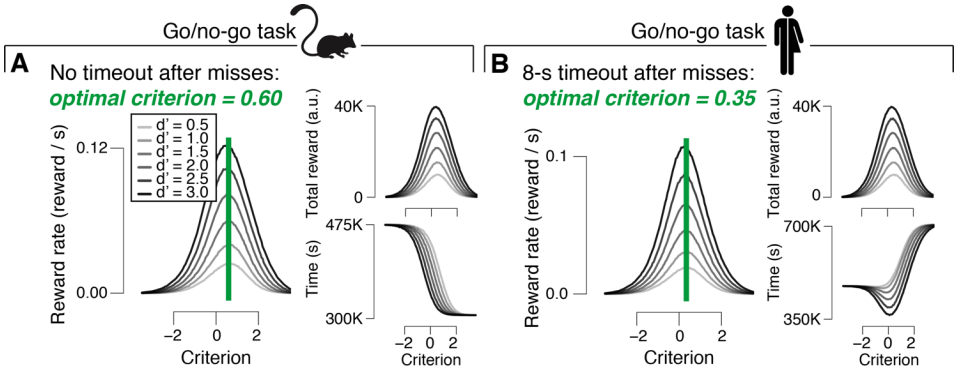

Figure S2. Optimality analysis of go/no-go task. (A) Results of simulation study of optimal bias in go/no-go task for mice (Materials and Methods). Optimality was defined as the level of choice bias (signal detection criterion) that maximized reward rate (# rewards / total time). Left: reward rate as a function of criterion, separately for six levels of sensitivity. Top and bottom, as left, but for total reward and elapsed time, respectively. **(B)** As A, but for the human version of the go/no-go task, with additional timeouts after misses.

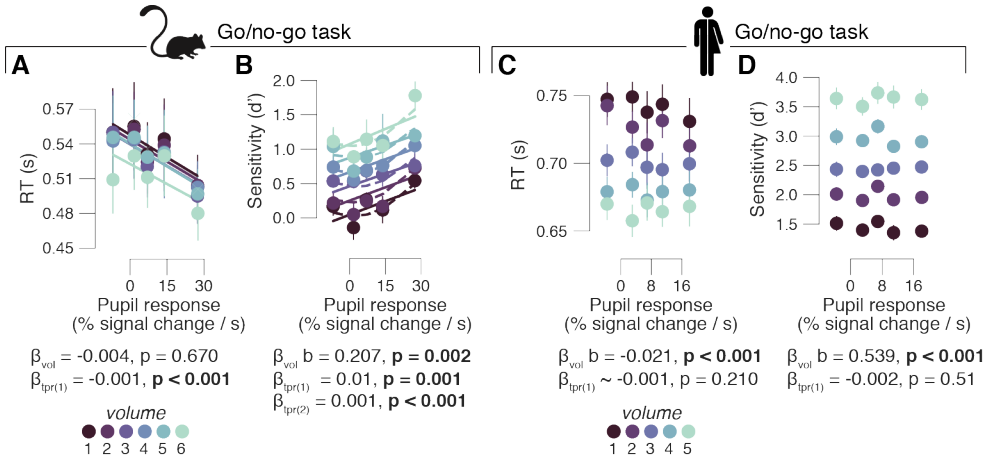

Figure S3. Relationship between pupil response and RT and perceptual sensitivity. (A) Relationship between median RT (dots, mean across subjects) and pupil response in mice, separately for the different signal strengths. Linear fits are plotted wherever the first-order fit was superior to the constant fit (see Materials and Methods). Quadratic fits were plotted (dashed lines) wherever the second-order fit was superior to first-order fit. Stats, mixed linear modeling. **(B)** As A, but for perceptual sensitivity (quantified by signal detection d'). **(C,D)** As A,B, but for humans. All panels: group average (N = 5; N = 20); error bars, s.e.m.

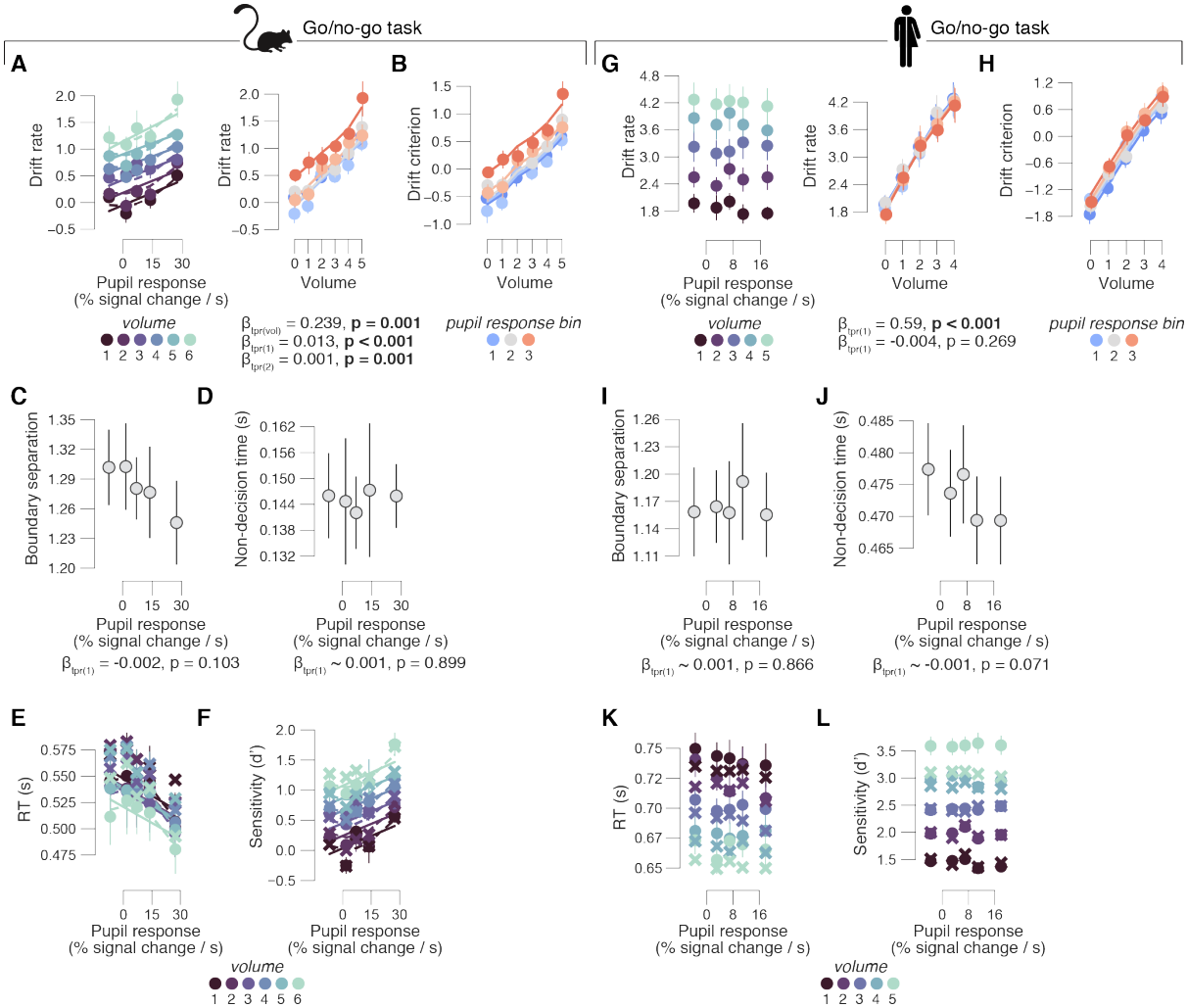

Figure S4 Pupil-dependent changes in computational model parameters during go/no-go task. **(A)** Left: relationship between drift rate and pupil responses in mice, separately for the different signal strengths. Right: as left, but with signal strength on the x-axis, and separate colors for the pupil-defined bins. Linear fits are plotted wherever the first-order fit was superior to the constant fit (Materials and Methods). Quadratic fits were plotted (dashed lines) wherever the second-order fit was superior to first-order fit. Stats, mixed linear modeling. **(B)** As A right, but for drift criterion. See main Fig. 3 for stats. **(C)** Relationship between boundary separation and pupil response in mice. Boundary separation was fixed across signal strengths. **(D)** As C, but for non-decision time. **(E)** Empirical (dots; as in Fig. S3A) and model-predicted (x-markers; Materials and Methods) RT as a function of pupil response in mice, separately for the different signal strengths. (**F**) As E, but for sensitivity. **(G-L)** As A-F, but for humans. All panels: group average (N = 5; N = 20); error bars, s.e.m.

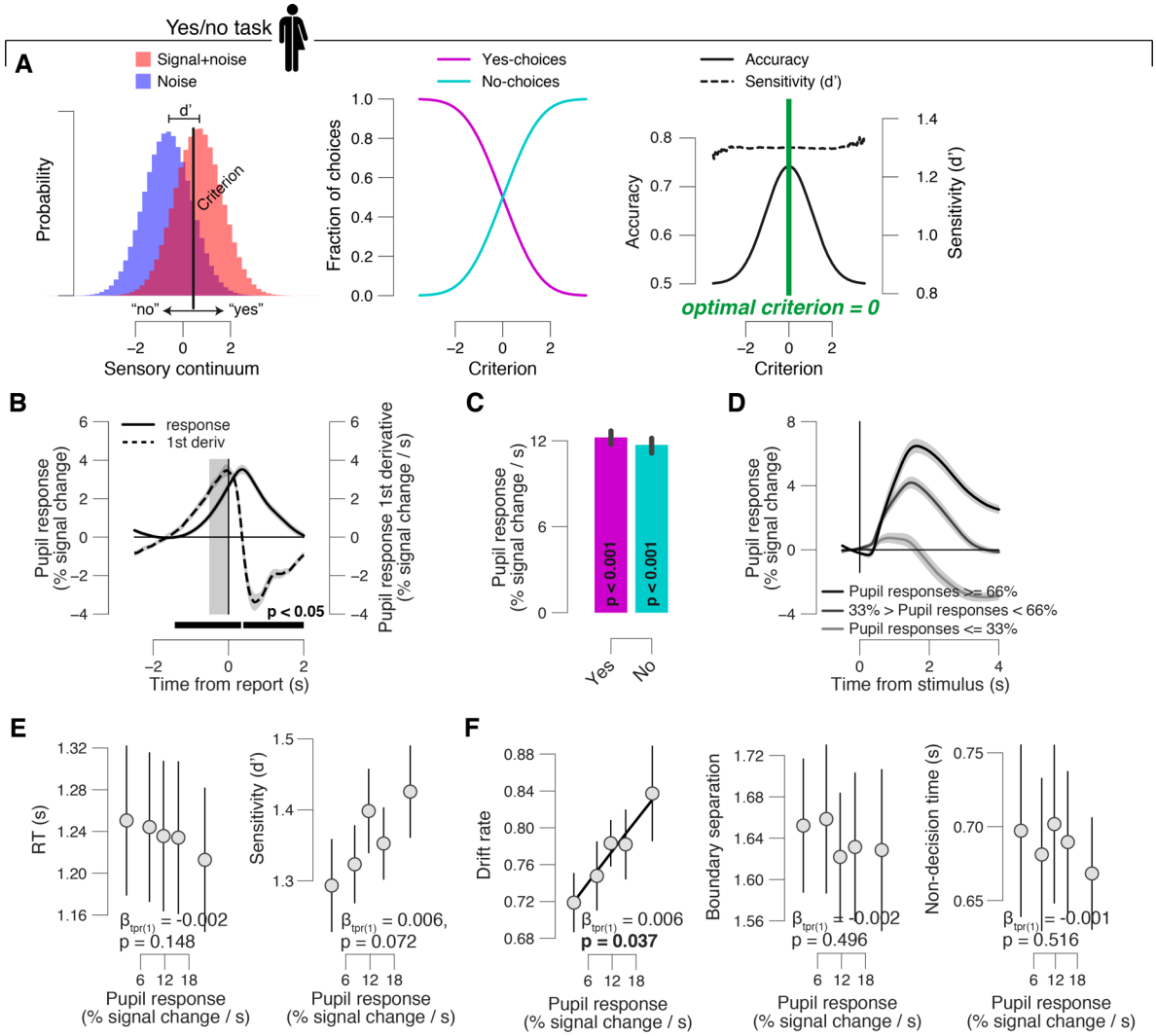

Figure S5. Optimality analysis of yes/no task. (A) Simulation study of optimal bias in yes/no (forced-choice) task (Materials and Methods). Optimality was defined as the level of choice bias (signal detection criterion) that maximized accuracy. Left: simulated internal sensory representations for signal+noise and noise trials. Signal detection theory assumes that internal sensory representations are normally distributed across trials (with the same standard deviation) and shifted between signal+noise and noise trials along the sensory continuum (in our case, volume of the pure sine-wave). Subject will base their signal detection judgement according to some criterion along the sensory continuum ("Criterion"). Middle: fraction of yes- and no-choices as a function of criterion (obtained after sliding the criterion along in the sensory continuum in the left panel). Right: as middle panel but for accuracy and perceptual sensitivity (d'). When signal+noise and noise trials are equally frequent, the optimal signal detection criterion is zero, because a neutral bias maximizes accuracy. **(B)** Task-evoked pupil response (solid line) and response derivative (dashed line). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. Stats, paired-samples t-test. **(C)** Task-evoked pupil responses sorted into yes- and no-choices. Stats, paired-samples t-test. **(D)** Overall pupil response time course for three pupil derivative defined bins. **(E)** Left: Relationship between RT and pupil response. Right: as left, but for perceptual sensitivity (quantified by signal detection d'). **(F)** As E, but for drift rate (left), boundary separation (middle) and non-decision time (right). All panels: group average (N = 24); shading or error bars, s.e.m.

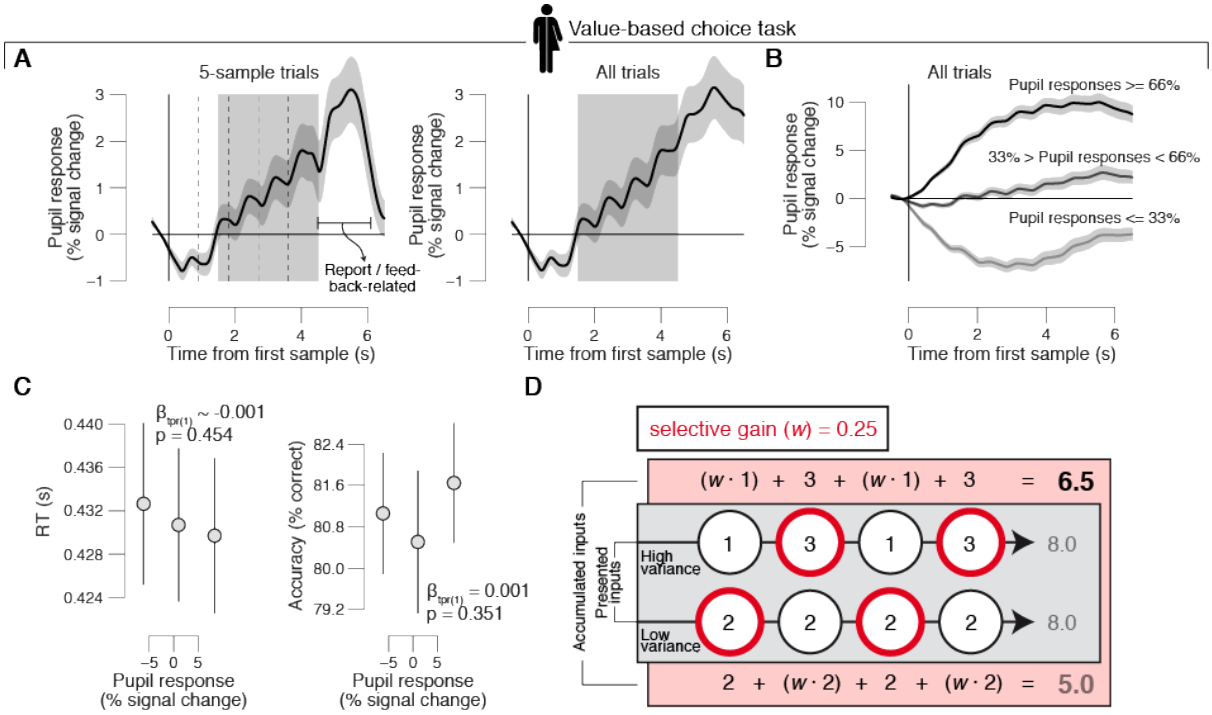

Figure S6. Pupil responses and pupil-dependent changes in behavior during value-based choice task. (A) Task-evoked pupil response time courses locked to the onset of the first pair of samples. Left: for the shortest trials (5 pairs of samples). Dashed vertical lines, onset of sample pairs 2-5. Right: for all trials (5–8 pairs of samples). Grey box, interval for computing single trial task-evoked pupil response measures (see Materials and Methods). **(B)** Overall pupil response time course for three overall pupil defined bins. **(C)** Left: relationship between RT and task-evoked pupil responses (3 bins). Right: as left, but for accuracy. Linear fits are plotted wherever the first-order fit was superior to the constant fit (see Materials and Methods). Quadratic fits were not superior to first-order fits. Stats, mixed linear modelling. **(D)** Toy-example of the selective integration model. Within the selective integration model, the pro-variance effect is attributed to the selective gain parameter *w*. Consider two value sequences with generative distributions of unequal variance. A losing sample in the high variance sequence (1-3-1-3) will more likely have a low value. Conversely, a losing sample in the low variance sequence (2-2-2-2) will more likely have a intermediate value. Multiplicatively downweighting a low value results in a smaller loss than multiplicatively downweighting a intemediate value (for a value of 1 and for $w = 0.25$ the loss is 0.75; for a value of 2 the respective loss is 1.5). Therefore, the accumulated value at the end of the trial will tend to be higher for the high variance sequence than for the low variance sequence; in this toy-example 6.5 vs. 5.0, respectively. All panels: group average (N = 32); error bars or shading, s.e.m.