Phasic arousal suppresses suboptimal decision biases in mice and humans

J. W. de Gee[1,2,3,4], K. Tsetsos[1], L. Schwabe[5], A.E. Urai[1,2,6], D. A. McCormick[7,8], M. J. McGinley[3,4,8*,#], T. H. Donner[1,2,9*]

[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]Department of Cognitive Psychology, Institute of Psychology, University of Hamburg, Germany; [6]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA; [7]Institute of Neuroscience, University of Oregon, OR, USA; [8]Department of Neuroscience, Yale University, New Haven, CT, USA; [9]Amsterdam Brain & Cognition, University of Amsterdam, Amsterdam, Netherlands.

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

Many difficult decisions are made by accumulating ambiguous evidence over time. The brain's arousal systems are rapidly activated during such decisions. How do these rapid ("phasic") boosts in arousal affect the decision process? Here, we have unveiled a general principle of the function of phasic arousal: suppressing suboptimal biases in evidence accumulation. We quantified phasic arousal as rapid dilations of the pupil. Pupil dilations predicted reduced biases in a range of decision-making tasks and different species. In a challenging sound-detection task, both mice and humans were less biased under high arousal. Similar bias suppression occurred when optimal biases were neutral, conservative or liberal, when evidence was accumulated from memory, and for risk-seeking biases in decisions entailing the accumulation of numerical values. In all cases, the smaller behavioral biases were explained by specific changes in evidence accumulation. Thus, phasic arousal calibrates a key computation during decision-making.

INTRODUCTION

The global arousal state of the brain changes from moment to moment (Aston-Jones & Cohen, 2005; McGinley, Vinck, et al., 2015). These global state changes are controlled in large part by modulatory neurotransmitters released from subcortical nuclei such as the noradrenergic locus coeruleus and the cholinergic basal forebrain. Release of these neuromodulators can profoundly change the operating mode of target cortical circuits (Aston-Jones & Cohen, 2005; Froemke, 2015; Harris & Thiele, 2011; S.-H. Lee & Dan, 2012; Pfeffer et al., 2018). These same arousal systems are phasically recruited during elementary

decisions, in relation to key computational variables such as uncertainty and surprise (Aston-Jones & Cohen, 2005; Bouret & Sara, 2005; Colizoli, de Gee, Urai, & Donner, 2018; Dayan & Yu, 2006; Krishnamurthy, Nassar, Sarode, & Gold, 2017; Lak, Nomoto, Keramati, Sakagami, & Kepecs, 2017; Nassar et al., 2012; Parikh, Kozak, Martinez, & Sarter, 2007; Urai, Braun, & Donner, 2017).

Phasic arousal during decision-making might play a key role in calibrating the biases that shape behavior and bound the rationality of judgment and decision-making in (Kahneman, 2011). Influential theoretical accounts propose that phasic arousal has an adaptive function that serves to optimize inference and choice behavior (Aston-Jones & Cohen, 2005; Dayan & Yu, 2006). Yet, the precise functional consequences of phasic arousal remain elusive, largely due to technical limitations in monitoring activity in these deep-brain structures during behavior. Here, we set out to resolve four outstanding issues pertaining to the adaptive function of phasic arousal.

First, little is known about the specific impact of phasic arousal on the transformation of decision-relevant evidence into a behavioral choice. Most decisions – including judgments about weak sensory patterns embedded in time-varying noise – are based on a protracted deliberation process (Shadlen & Kiani, 2013). This process seems to be implemented by a distributed brain network: association cortex accumulates input signals ("evidence samples") over time, into a decision variable, and motor regions translate the decision into a behavioral act (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Shadlen & Kiani, 2013; Siegel, Engel, & Donner, 2011; Wang, 2008). Since arousal shapes the state of all these brain regions, phasic arousal might alter the encoding of the evidence, the accumulation of the evidence, the implementation of the motor act, or all of the above. One influential account proposes that phasic arousal specifically speeds up the translation of a choice into the resulting motor act (Aston-Jones & Cohen, 2005). We asked whether and how phasic arousal might also shape the preceding evidence accumulation.

Second, while brainstem arousal systems are homologously organized across mammals (Amaral & Sinnamon, 1977; Berridge & Waterhouse, 2003), it is not clear whether arousal systems are recruited in the same circumstances across species. More specifically, it is not clear whether the computations underlying decision formation under uncertainty are affected by arousal signals in a manner that is consistent across species. Rodents (rats) and humans seem to accumulate perceptual evidence in a similar fashion (Brunton, Botvinick, & Brody, 2013). But is the shaping of this computation by phasic arousal also governed by a general principle?

Third, elementary perceptual decisions provide an established laboratory model of evidence accumulation (Shadlen & Kiani, 2013), but many important real-life decisions (e.g. which stock to buy, or which career to pursue) are based on the accumulation of non-sensory signals, such as those gathered from memory (Shadlen & Shohamy, 2016), or from abstract (e.g. numerical) quantities (Tsetsos, Chater,

& Usher, 2012). Does phasic arousal have the same impact on decisions requiring the accumulation of perceptual versus higher-level evidence?

Fourth, biases in choice behavior can either be adaptive or maladaptive, depending on the statistics of the environment (Green & Swets, 1966). For example, in many laboratory perceptual choice tasks, stimuli are equally likely to occur, so any bias needs to be suppressed to optimize behavior. But when a certain target is more (or less) likely to occur, choice should be biased towards (or away) from that target choice. It is not known whether phasic arousal flexibly affects choice biases based on stimulus statistics, or, for example, if phasic arousal makes animals and humans uniformly more liberal in their decisions.

We approached these issues by means of a cross-species, integrated behavioral and computational approach. We combined pupillometry, behavioral experiments and modeling, in both humans and mice, and studied humans in a variety of behavioral contexts. Pupil dilation is a reliable peripheral proxy of several established markers of cortical arousal state (McGinley, David, & McCormick, 2015; Reimer et al., 2014; Vinck, Batista-Brito, Knoblich, & Cardin, 2015). Our results have revealed a general principle regarding the function of phasic arousal in decision-making: suppressing biases in evidence accumulation. Thus, the protracted deliberation underlying decisions (Shadlen & Kiani, 2013) is shaped by task-evoked neuromodulatory responses.

RESULTS

Humans and mice performed the same simple perceptual decision (auditory go/no-go detection). In addition, humans performed a forced-choice decision task based on the same auditory evidence under systematic manipulations of target probabilities, a memory-based decision task, and a basic laboratory task model of value-based stock market decisions.

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; Materials and Methods). Subjects searched for a signal (pure tone) embedded in a sequence of discrete, but dynamic, noise tokens. Because signals were embedded in fluctuating noise, detection performance could be maximized by accumulating the sensory evidence over time. To indicate a yes-choice, mice licked for sugar water reward and human subjects pressed a button. The loudness was manipulated by varying the sound level (volume) of the tone, while keeping the noise level constant. As expected, in both species, reaction times (RT) parametrically decreased with loudness (Fig. 1B) and signal detection-theoretic sensitivity (d'; Materials and Methods) parametrically increased (Fig. 1C). Humans responded overall a little slower than mice (Fig. 1B).

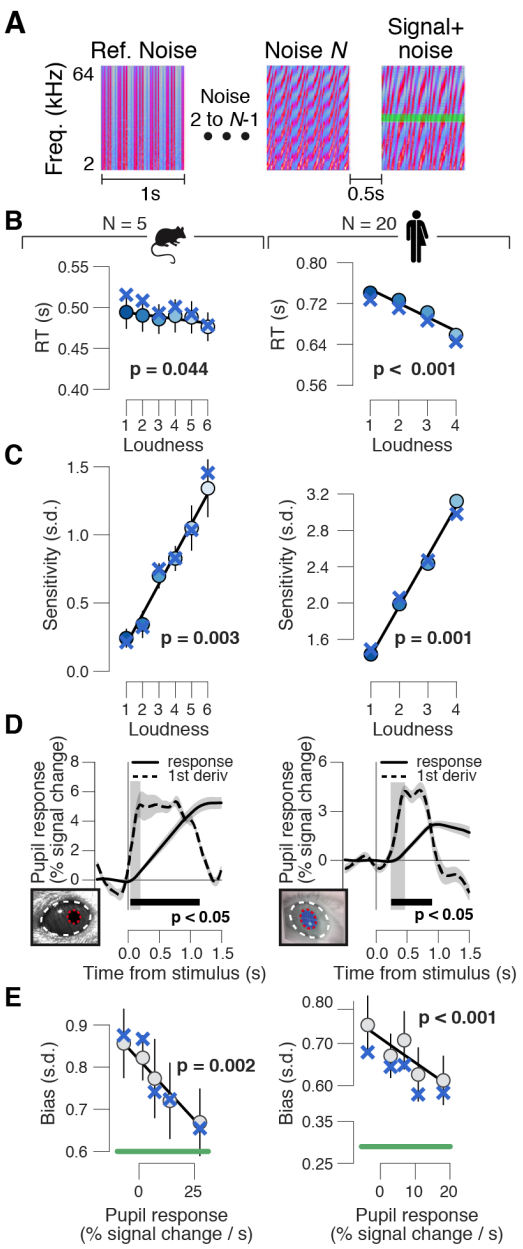

Figure 1. High phasic arousal is associated with reduced perceptual choice bias. (A) Auditory go/no-go tone-in-noise detection task. Schematic sequence of stimuli during a trial. Subjects were trained to respond to a weak signal (pure tone) in noise and withheld a response for noise-only sound tokens. Each sound was treated as a separate decision (Materials and Methods). **(B)** Relationship between reaction time and loudness in mice (left) and humans (right). 'X' symbols are predictions from best fitting variant of drift diffusion model (Materials and Methods); stats are from mixed linear modeling (Materials and Methods). **(C)** As panel B, but for sensitivity (quantified by signal detection d'; Materials and Methods**). (D)** Task-evoked pupil response (solid line) and response derivative (dashed line) in mice (left) and humans (right). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative; stats, cluster-corrected one-sample t-test. **(E)** Relationship between overall perceptual choice bias (Materials and Methods) and task-evoked pupil response in mice (left) and humans (right). Linear fits were plotted if first-order fit was superior to constant fit; quadratic fits were not superior to first-order fits (Materials and Methods). 'X' symbols are predictions from best fitting variant of drift diffusion model (Materials and Methods); green line, optimal bias (see Fig. S1C,D,J); stats, mixed linear modeling. Panels B-E: group average (N = 5; N = 20); shading or error bars, s.e.m.

To track phasic arousal, we measured the rising slope of the pupil, immediately after each sound onset. We choose this measure for three reasons: (i) for its temporal precision in tracking arousal during fast-paced tasks (Fig. 1D), (ii) to eliminate contamination by movements (licks and button-presses) (de Gee, Knapen, & Donner, 2014; Hupé, Lamirel, & Lorenceau, 2009) (Materials and Methods), and (iii) to most specifically track noradrenergic activity (Reimer et al., 2016). The timing of stimuli was predictable, so subjects could tightly align a phasic arousal response to the next sound onset. Indeed, pupil responses occurred from 40 ms after sound onset in mice (Fig. 1D), and from 240 ms after sound onset in humans (Fig. 1D). The shorter pupil response latencies in mice compared to humans might be due to their smaller eye and brain size. Pupil responses occurred also on trials without a behavioral

response (Fig. S1F,L), consistent with other observations (C. R. Lee & Margolis, 2016; Schriver, Bagdasarov, & Wang, 2018).

In both mice and humans, we found a consistent relationship between the early, task-evoked pupil response and decision outcome. Because loudness was drawn pseudo-randomly on each trial, subjects had different d' values for each loudness (Fig. 1C) but could set only one decision criterion (or bias set point) against which to compare sensory evidence. Therefore, using signal detection theory, we computed an overall perceptual choice bias across loudness (Fig. S1C; Materials and Methods). We found that both mice and humans had an overall conservative perceptual choice bias, preferably failing to respond that they perceived the tone (Fig. 1E). This conservative bias was maladaptive, reducing the fraction of correct/rewarded choices below what could be achieved at a given perceptual sensitivity (Fig. S1D,J). In both species this maladaptive bias was suppressed on trials with large pupil responses (Fig. 1E). The same was true when computing overall bias as the average signal detection theoretic criterion across loudness (Fig. S1I,O). Phasic pupil responses exhibited a less consistent relationship to perceptual sensitivity and RT (Fig. S1H,N).

Previous work has associated baseline, pre-stimulus arousal state with non-monotonic (inverted U-shape) effects on decision performance (Aston-Jones & Cohen, 2005; Yerkes & Dodson, 1908), even in the same mice dataset analyzed here (McGinley, David, et al., 2015). By contrast, we here found that the dominant predictive effect of pupil-linked phasic arousal was a monotonic (linear) reduction of bias (Fig. 1E, solid lines; Materials and Methods), pointing to distinct functional roles of tonic and phasic arousal (Discussion).

Phasic arousal predicts a reduction of evidence accumulation bias

Our analyses of overt behavior revealed that pupil-linked phasic arousal was associated with a largely monotonic reduction of maladaptive choice bias in mice and humans. Fitting decision-making models enabled us to gain deeper insight into how the decision process was affected by phasic pupil-linked arousal. We fitted the drift diffusion model (Fig. S2A), which belongs to a class of bounded accumulation models of decision-making (Bogacz et al., 2006; Brody & Hanks, 2016; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008) that describe the accumulation of noisy sensory evidence in a 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, Huang-Pollock, & McKoon, 2016; Ratcliff & McKoon, 2008). We used the diffusion model to quantify effects of pupil-linked arousal the following components of the decision process: the starting point of evidence accumulation, the evidence accumulation itself (the mean drift rate and an evidence-dependent bias in the drift, henceforth called "drift bias"), boundary separation

(implementing speed-accuracy tradeoff) and the so-called non-decision time (the speed of pre-decisional evidence encoding and post-decisional translation of choice into motor response).

In order to assess phasic arousal-dependent changes in all of these parameters, we fit a model in which all parameters (except for starting point) were free to vary with pupil response amplitude. The absence of RTs for no-responses in the go/no-go datasets forced us to fix either starting point or drift bias as function of pupil. We chose to fix starting point because formal model comparison favored this over fixing drift bias in all subjects of both species (Materials and Methods). Indeed, in all other datasets analyzed in this study, presented below, we found that the pupil-linked bias suppression in overt behavior was specifically due to a shift in drift bias, not starting point.

The model accounted well for the overall behavior in the go/no-go task. First, as expected, drift rate increased with loudness, reflecting the subjects' ability to accumulate strong sensory evidence more efficiently (Fig. S2F,I). Second, the fitted parameters accurately predicted overall RT and sensitivity (Fig. 1B,C).

In both species, we found a positive linear relationship between pupil responses and drift bias (Fig. 2C). The fitted parameters accurately predicted overall perceptual choice bias, and its pupil response predicted shift (blue 'X' markers in Fig. 1E). Specifically, in both species, the starting point was biased towards no-go irrespective of pupil response (Fig. S2E,H). Thus, overcoming this conservative choice bias required increasing their drift bias. Such an increase in drift bias occurred on trials with large pupil responses (Fig. 2C).

Phasic arousal had no, or less consistent, effects on the other model parameters. There was no consistent monotonic effect of pupil response on boundary separation, drift rate or non-decision time in either species (mice: $p = 0.722$, $p = 0.073$ and $p = 0.269$, respectively; humans: $p = 0.484$, $p = 0.15$ and $p = 0.132$, respectively; Fig. 2). Without collapsing across loudness we observed a positive (negative) relationship between drift rate pupil response in mice (humans) (Fig. S2D,G), and a negative relationship between pupil response and non-decision time in both mice and humans (Fig. S2D,G). These effects on drift rate and non-decision time were however not consistent across the variety of behavioral contexts considered here. These results suggest that the dominant impact of phasic arousal in this task was remarkably specific: optimizing the evidence accumulation process by suppressing a bias in the drift.

Perceptual choice variability has been attributed to evidence accumulation noise, rather systematic accumulation biases, under the assumption that biases will remain constant across trials (Drugowitsch, Wyart, Devauchelle, & Koechlin, 2016). Instead, our results show that accumulation biases vary dynamically across trials as a function of phasic arousal. This indicates the resulting choice variations should appear as random trial-by-trial variability in evidence accumulation when ignoring phasic arousal. We found that this was the case in our data (Fig. 2E). We simulated RT distributions from

two conditions that differed according to the fitted drift bias estimates in the lowest and highest pupil-defined bin of each individual (Materials and Methods). The diffusion model accounts for trial-to-trial accumulation "noise" with the drift rate variability parameter (Bogacz et al., 2006; Ratcliff & McKoon, 2008). Indeed, when fitting the model to these simulated RT distributions, drift rate variability was accurately recovered when drift bias could vary with condition but was significantly overestimated when drift bias was fixed (Fig. 2E). This analysis is agnostic about the source of trial-by-trial variations in phasic arousal, which was not under experimental control in the present study (but see (Colizoli et al., 2018; Nassar et al., 2012; Urai et al., 2017). But the results clearly show that a significant fraction of choice variability does not originate from noise within the evidence accumulation machinery, but rather from the neural systems that govern arousal.

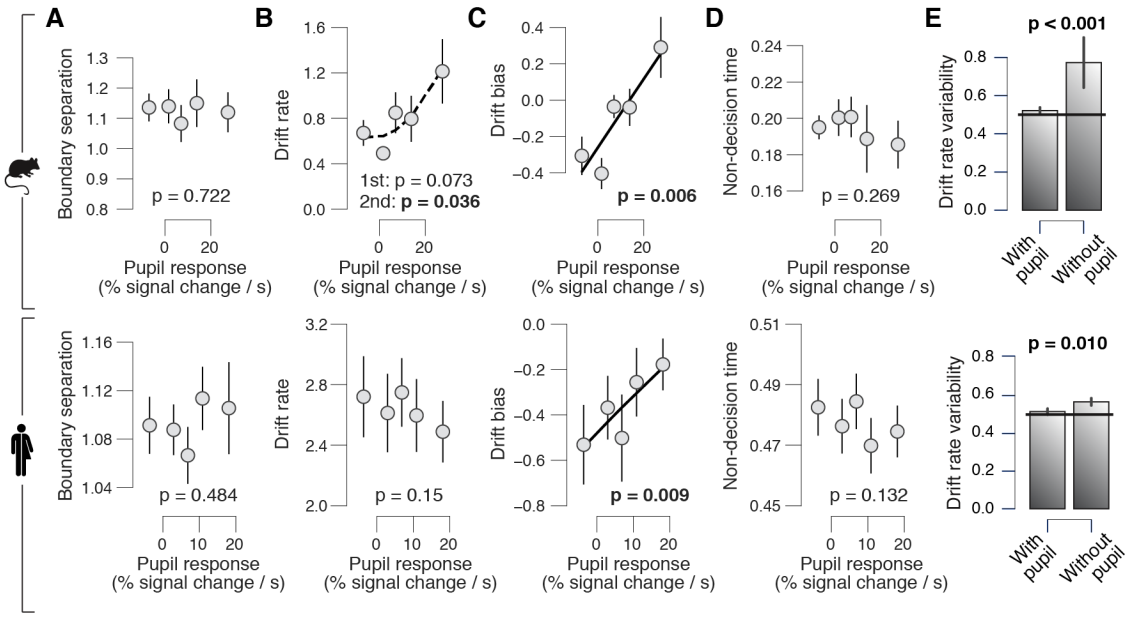

Figure 2. Phasic arousal accounts for substantial variability in sensory evidence accumulation. (A) Relationship between boundary separation estimates and task-evoked pupil response in mice (top) and humans (bottom), collapsed across loudness. See Fig. S2 for parameter estimates separately per loudness. Linear fits are plotted wherever the first-order fit was superior to the constant fit; quadratic fits were plotted (dashed lines) wherever the second-order fit was superior to first-order fit. Stats, mixed linear modeling. **(B-D)** As A, but for drift rate, drift bias and non-decision time estimates, respectively. **(E)** Recovered drift rate variability for models with and without pupil predicted shift in drift bias. The model was fit to simulated RT distributions from two conditions that differed according to the fitted drift bias estimates in the lowest and highest pupil-defined bin of each individual (Materials and Methods). Black line, true drift rate variability. Stats, paired-samples t-test. All panels: group average (N = 5; N = 20); error bars, s.e.m.

Phasic arousal predicts accumulation bias suppression in forced choice version of the detection task

Are the above results specific to the go/no-go protocol used to study the auditory detection decision? The central input to the pupil contains a sustained component during evidence accumulation, followed by a transient at the motor response (de Gee et al., 2017, 2014; Hupé et al., 2009; Murphy, Boonstra, & Nieuwenhuis, 2016). The sustained component might entail motor preparatory activity (Donner et al, 2009). Thus, a concern about the go/no-go task might be that these components (motor preparatory activity and transient activity at lick / button-press) could have contributed to the pupil response amplitudes on go-trials but not (or less so) on no-go-trials, which in turn could explain the relationship between pupil responses and choice bias. We reduced contamination by the transient motor-related component by focusing on the initial (early) pupil dilation (Materials and Methods). However, it is possible that this did not fully correct for the asymmetry between go- and no-go trials in motor preparatory activity.

We asked human participants (N = 24, 18 from the above go no-go experiment) to perform a forced-choice (yes/no) version of the above auditory detection task, based on the same type of auditory evidence. In this task, motor responses (and associated preparatory activity) were balanced across yes- and no-choices (Fig. 3A). Consistent with our go/no-go results in mice and humans, we observed that the pupil response predicted a monotonic suppression of maladaptive perceptual choice bias (Fig. 3B) again, pushing behavior to a more optimal regime (Fig. S3A). Pupil response amplitudes in the go/no-go and yes/no tasks were correlated across eighteen human subjects who participated in both experiments (Fig. S3D). This was true for yes-choices and no-choices. Therefore, the suppression of choice bias in our results does not reflect motor preparation.

Diffusion modeling again revealed a performance-optimizing effect of pupil response on evidence accumulation: a reduction in drift bias, here accompanied by an increase in mean drift rate (Fig. 3C). With respect to pre- or post-decisional parameters, there was a non-monotonic effect on starting point ($p = 0.013$) and again no effect on boundary separation and non-decision time ($p = 0.327$ and $p = 0.722$, respectively). Critically, the pupil-linked changes in drift bias, but not the changes in starting point, strongly correlated with the individual reductions in decision bias as measured by SDT in Fig. 3B (squared multiple correlation $R^2 = 0.952$; drift bias: beta = -1.01, $p < 0.001$; starting point: beta = -0.10, $p = 0.219$). Thus, only the changes of drift bias explained the performance-optimizing reductions in decision bias.

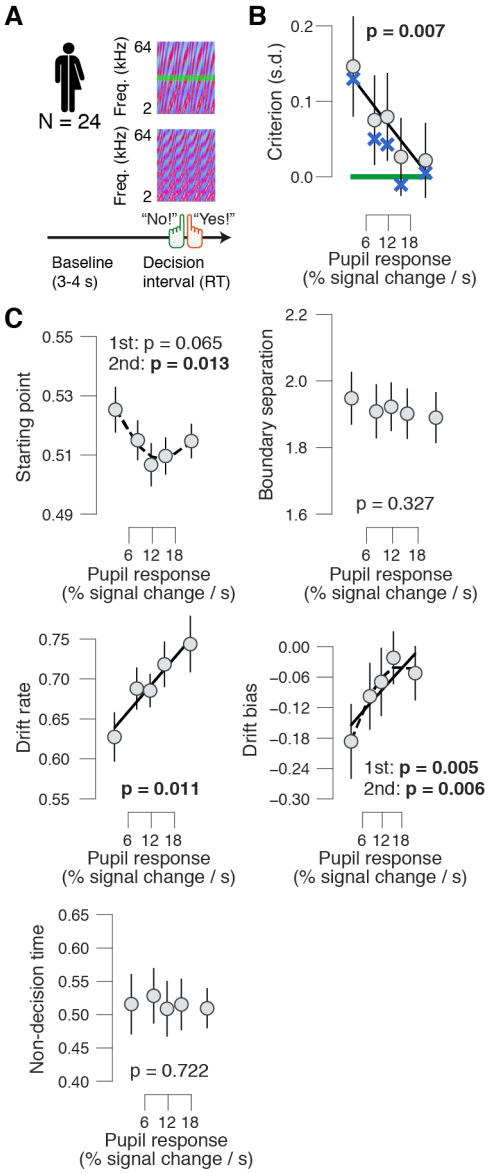

Figure 3. High phasic arousal associates with reduced perceptual evidence accumulation bias during yes/no decisions. (A) Auditory yes/no (forced choice) tone-in-noise detection task. Schematic sequence of events during a trial. Subjects reported the presence or absence of a faint signal (pure sine wave) embedded in noise (Materials and Methods). **(B)** Relationship between perceptual 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 were plotted (dashed lines) wherever the second-order fit was superior to first-order fit. Green line, optimal criterion (see Fig S3A); 'X' symbols are predictions from the drift diffusion model; stats, mixed linear modeling; group average (N = 24); error bars, s.e.m. **(C)** As B, but for relationship between drift diffusion model parameters and task-evoked pupil response.

Phasic arousal predicts a reduction of conservative and liberal perceptual accumulation biases

The majority of mice and humans in the above go/no-go and yes/no tasks exhibited a conservative tendency towards choosing "no." This conservative bias was suboptimal in these tasks, and so the pupil-linked increase in liberal decision-making (accumulation towards "yes") improved performance (see Figs. 1 and 3). However, when targets are rare (or false alarms heavily penalized), a conservative bias is optimal, and conversely for frequent targets (Green & Swets, 1966). Thus, for phasic arousal to be adaptive, its effects on accumulation biases should be flexibly adjustable to stimulus statistics; promoting liberal decision-making in a stereotypical fashion would not be adaptive.

To assess whether pupil-linked modulation of accumulation biases adapts to the stimulus statistics, we asked a new group of observers to perform the same auditory yes/no task, but with rare (P(Signal)=0.3) or frequent (P(Signal)=0.7) occurrence of targets (Material and Methods). As expected, subjects developed a conservative bias in the rare target condition and a liberal bias in the frequent target condition (Fig. 4A). We used three pupil-defined bins because there were fewer critical trials per individual (less than 500) than in the previous data sets (more than 500; Materials and Methods). Critically, pupil response now predicted changes in choice biases of opposite sign in the two conditions (Fig. 4B). In both conditions, increased pupil response was associated with a tendency towards neutral bias. Again, the effect of pupil-linked arousal on choice biases was mediated by shifts in accumulation biases (Fig. 4C). There was an effect of pupil-linked arousal on starting point too, but in the opposite direction as the choice bias shift (Fig. 4C). Again the pupil-linked changes in drift bias, but less so the changes in starting point, correlated with the individual reductions in decision bias as measured by SDT in the rare condition (squared multiple correlation $R^2 = 0.959$; drift bias: beta = -0.97, $p < 0.001$; starting point: beta = -0.07, $p = 0.039$) as well as in the frequent condition (squared multiple correlation $R^2 = 0.997$; drift bias: beta = -1.08, $p < 0.001$; starting point: beta = 0.29, $p = 0.024$). Thus, again only the changes of drift bias explained the reductions in decision bias.

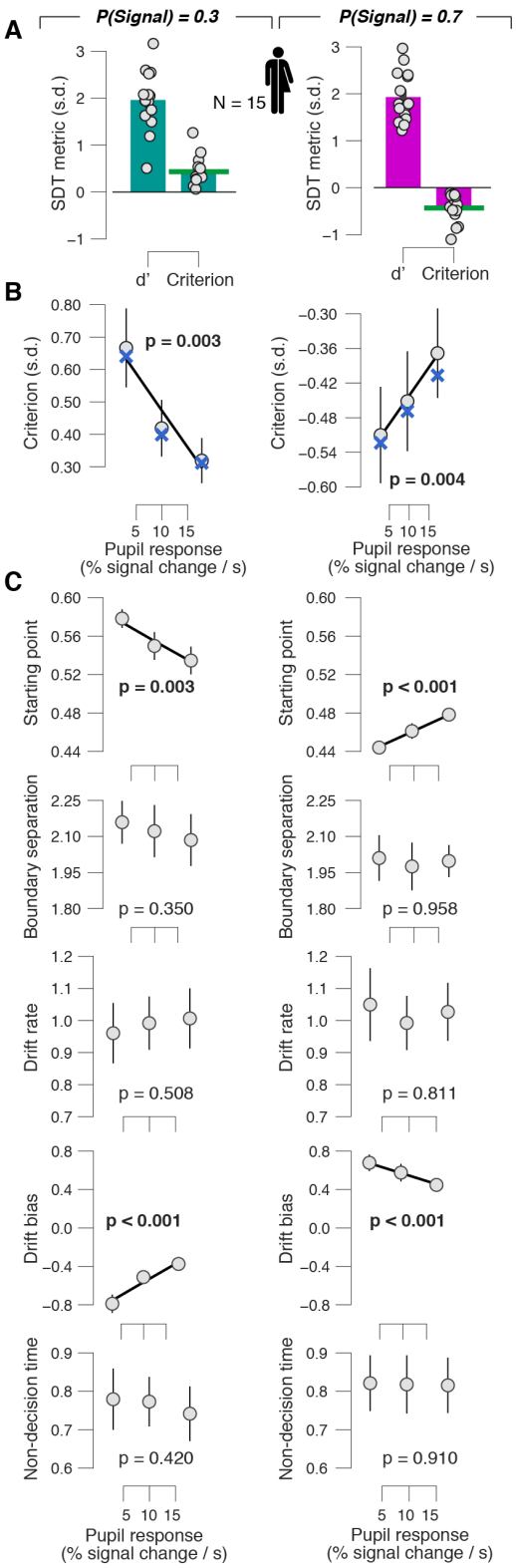

Figure 4. Phasic arousal flexibly reduces both conservative and liberal accumulation biases. (A) Overall sensitivity and choice bias (quantified by signal detection d' and criterion, respectively) in the rare signal condition (P(Signal) = 0.3; left) and frequent condition (P(Signal) = 0.7; right). Green line, optimal bias. See Fig. S4A,B for task-evoked pupil response time course. **(B)** Relationship between choice bias and task-evoked pupil response in the rare condition (left) and frequent condition (right). Linear fits were plotted if first-order fit was superior to constant fit; quadratic fits were not superior to first-order fits. 'X' symbols are predictions from the drift diffusion model; stats, mixed linear modeling. **(C)** As B, for relationship between drift diffusion model parameters and task-evoked pupil response. All panels: group average (N = 15); error bars, s.e.m.

Phasic arousal predicts a reduction of conservative and liberal memory accumulation biases

In many important real-life decisions, the evidence feeding into the decision process cannot be sampled from the current sensory environment, but rather needs to be drawn from memory. It has been proposed that memory-based decisions follow the same sequential sample principle established for perceptual decisions, whereby the "samples" accumulated into the decision variable are drawn from memory (Ratcliff, 1978; Shadlen & Shohamy, 2016). We next assessed whether the suppression of accumulation biases identified for perceptual decisions above generalized to memory-based decisions.

To this end, we modeled the impact of pupil-linked phasic arousal on choice behavior in a yes/no recognition memory task (Fig. 5A; Materials and Methods). Subjects were instructed to memorize 150 pictures (intentional encoding) and to evaluate how emotional each picture was on a 4-point scale from 0 ("neutral") to 3 ("very negative"). Twenty-four hours post encoding, subjects saw all pictures that were presented on the first day and an equal number of novel pictures in randomized order, and indicated for each item whether it had been presented before ("yes – old") or not ("no – new"). Data from an analogous task have been successfully fitted with the diffusion model (Bowen, Spaniol, Patel, & Voss, 2016), indicating that the arousal component of the images specifically alters accumulation bias (called "memory bias"). Here, we show the impact of trial-to-trial variations in phasic arousal, as measured by pupil responses, factoring out variations in the stimulus material (Materials and Methods).

The large sample size (N=54) in this data set afforded another critical test of the adaptive nature of the pupil-linked arousal effect. We observed a robust relationship between subjects' overall choice bias, and the pupil predicted shift in that bias: those subjects with the strongest biases, liberal or conservative, exhibited the strongest pupil-predicted shift towards a neutral (optimal) bias (Fig. 5B). Indeed, the group-average choice bias (signal detection theoretic criterion; sign-flipped for overall liberal subjects) was significantly reduced towards 0 (optimal) on large pupil response trials (Fig. 5C). Again, this effect of pupil-linked arousal on choice bias was mimicked by corresponding changes in accumulation bias (Fig. 5D, Fig. S5E,F) and not by changes in starting point (Fig. S5E,F) (difference in correlation: $\Delta r = -0.599$, $p < 0.001$). Note that lower criterion values indicate more liberal behavior, and lower drift bias values more conservative behavior, which explains the correlations of opposite sign in Fig. 5 panels B and D. The pupil response further predicted an increase in drift rate ($p = 0.01$, an increase (i.e. lengthening) on non-decision time, and no changes in starting point ($p = 161$) or boundary separation ($p = 0.089$) (Fig. S5E). Taken together, we conclude that phasic arousal reduces choice biases, irrespective of sign, and thus can shift both conservative and liberal biases towards optimality.

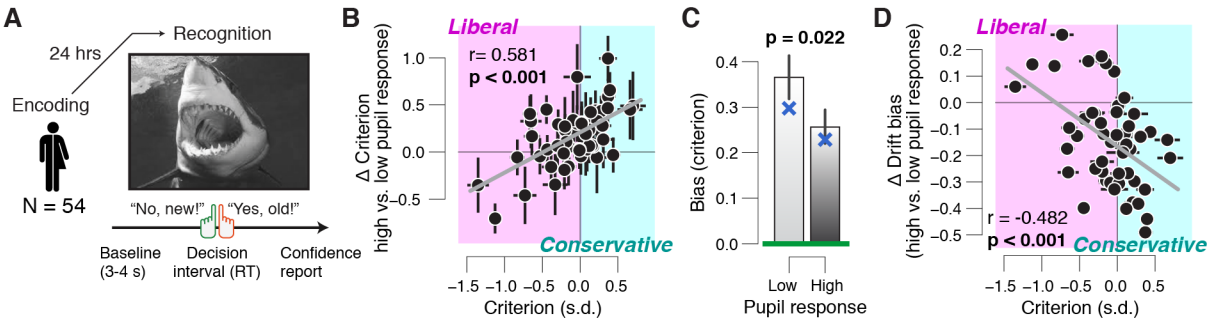

Figure 5. Phasic arousal predicts a reduction of memory accumulation biases. (A) A Picture yes/no (forced-choice) recognition task. Schematic sequence of events during a trial. Subjects judged whether they had seen pictures twenty-four hours previously during an encoding task (Materials and Methods). See Fig. S5A,B for task-evoked pupil response time course. **(B)** Individual pupil predicted shift in choice bias (quantified by signal detection criterion), plotted against individual's overall choice bias. Data points, individual subjects. Stats, Pearson's correlation. Error bars, 60% confidence intervals (bootstrap) **(C)** Choice bias (sign-flipped for overall liberal subjects) for low and high pupil response bins. Green line, optimal bias; 'X' symbols are predictions from the drift diffusion model; stats, paired-samples t-test; group average (N = 54); error bars, s.e.m. **(D)** As B, but pupil predicted shift in drift bias, plotted against individual's overall choice bias.

Arousal-linked bias reduction generalizes to high-level decisions

Does the arousal-related suppression of accumulation bias generalize to more high-level forms of biases identified in behavioral economics? Systematic deviations of human decision-making from rational choice are ubiquitous in value-based decision-making and higher-level reasoning (Tsetsos et al., 2012, 2016; Tversky & Kahneman, 1974). One form of such biases is risk-seeking, the tendency to choose options with large uncertainty about their outcome.

To study the impact of phasic arousal on risk-seeking, we used a task that probes decisions based on varying numerical information, akin to deciding which of two fluctuating stock options had the higher returns in the past year (Tsetsos et al., 2012, 2016). 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 (in this case numbers), it is amenable to the sequential sampling modeling approach we applied to perceptual decision-making above. We 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 Materials and Methods). Risk preference was quantified as a "pro-variance bias": the fraction of high-variance choices pooled across both trial types (Materials and Methods). As in previous work (Tsetsos et al., 2012, 2016), 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 used three pupil-defined bins because of the relatively few critical trials per individual (less than 500; Materials and Methods).

We again used sequential sampling modeling to pinpoint the source of the pupil-linked pro-variance bias suppression. Previous work on this task has uncovered specific deviations in the evidence accumulation process from the one at play in standard perceptual choice tasks (Tsetsos et al., 2012). We thus first compared the ability of four established decision-making models to account for behavior in the task (see Materials and Methods for mathematical descriptions): the drift diffusion model (DDM) that well accounted for all the previous data sets; the leaky accumulator model (LAM; (Bogacz et al., 2006)); the leaky competing accumulator model (LCAM; (Usher & McClelland, 2001)); and the leaky selective accumulator model (LSAM; (Tsetsos et al., 2016)). All models entail the perfect (DDM) or leaky (all other models) accumulation of the presented pairs of numbers across the trial. In the LSAM, accumulation is biased by a so-called selective gain parameter that prioritizes (i.e. assigns larger weight to) to the number that is higher on a given pair, giving rise to the pro-variance bias observed in choice behavior (Tsetsos et al., 2016).

Two separate criteria both favored the LSAM. First, group average BICs were lowest for the LSAM (242.30) compared to the DDM (248.48), LAM (243.64) and LCAM (248.73). BIC compares models based on their maximized log-likelihood value, while penalizing for the number of parameters (Schwarz, 1978). Lower BIC values indicate a model that better explained the data, whereby BIC differences of 10 indicate a decisively better fit (Schwarz, 1978). Second, and more importantly, only the LSAM was able to jointly account for two diagnostic behavioral signatures: (i) recency bias, a tendency to rely more on recent than on early samples of evidence (Fig. 6E; Fig. S6D; Materials and Methods); and (ii) the pro-variance bias (Fig. 6C). The DDM, by assuming perfect accumulation, could not account for recency or pro-variance biases (Fig. 6E,F). The LAM and the LCAM included leak terms and could therefore account for the recency bias; however, both models failed to capture the pro-variance bias (but see Fig. S6E). Only LSAM could account for both features of behavior (Fig. 6D–F).

Having established the LSAM as the best-fitting model for this task, we then used the LSAM fits to evaluate the effects of phasic arousal on evidence accumulation. Consistent with previous studies, the selective gain parameter was larger than 0 (Fig. 6G), 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. 6G), mediating a reduction in the pro-variance bias. We did not find robust evidence for pupil-predicted changes in other model parameters such as leak (controlling the time-constant of accumulation) or noise. 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) if noise corrupts the accumulation process downstream the of 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 a group average). The stronger the pupil responses, the closer to optimal was 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 in a high-level task was associated with more optimal decision-making.

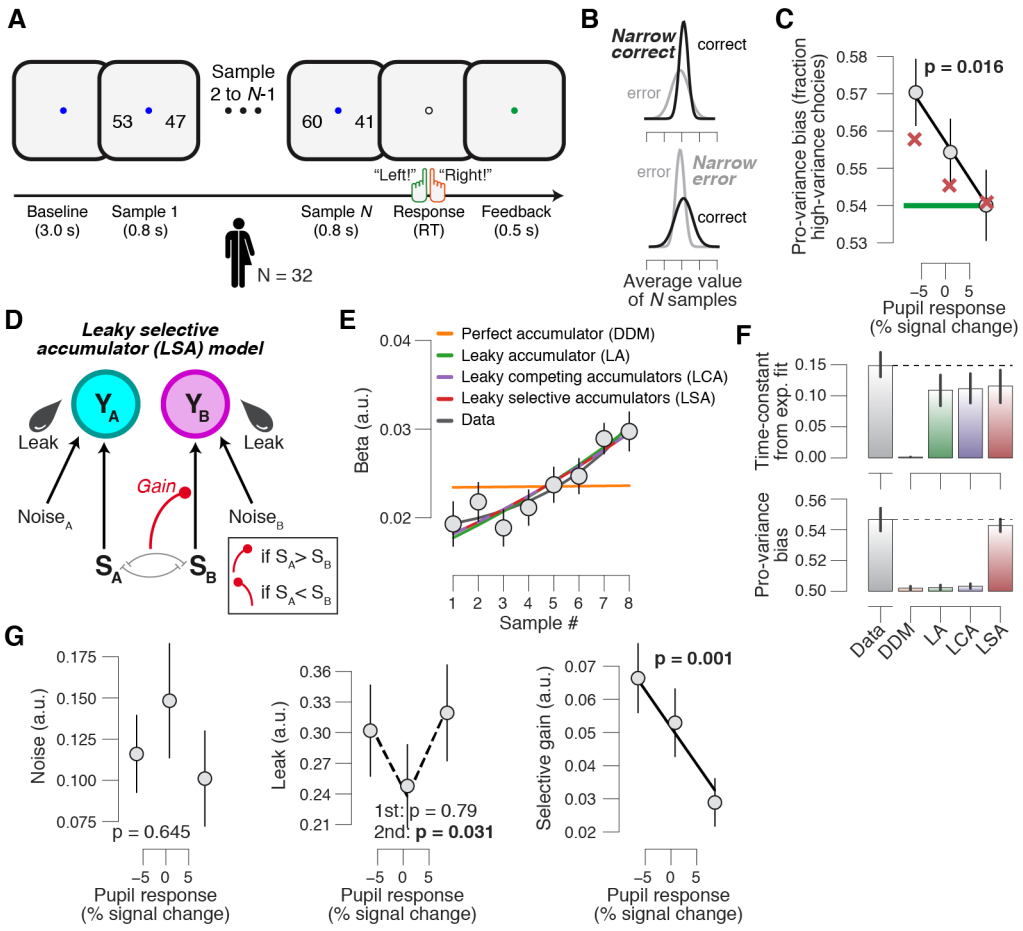

Figure 6. Phasic arousal predicts reduced risk seeking bias in a value-based choice task. (A) Value-based choice task. Schematic sequence of events during a trial. Subjects integrated two value sequences (5–8 pairs) and judged which sequence had on average the higher value. **(B)** On "narrow-correct" 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. 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. Green line, optimal fraction of high-variance choices; 'X' symbols are predictions from the drift diffusion model; stats, mixed linear modelling. **(D)** Schematic of the leaky selective accumulator model (LSAM; 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". **(E)** Psychophysical kernels, indicating the effect of each pair of numbers on observers' choice (plotted only for trials with 8 pairs; see Fig. S6D for all trial durations). Lines, exponential fits to data (grey) or model predictions (colors). **(F)** Time constant from exponential fit in E (top) and pro-variance bias (bottom) for the data (grey) and model predictions (colors). **(G)** As C, but for the LSAM parameters noise (left) and leak (middle) and selective gain (right). All panels: group average (N = 32); error bars, s.e.m.

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

One 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, Nieuwenhuis, Jepma, & Cohen, 2010), which likely causes floor and ceiling effects on pupil size due to eye muscle geometry and light conditions. Inherited correlation with baseline pupil 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: $r = 0.014$, ± 0.028 s.e.m.; humans: $r = -0.037$, ± 0.017 s.e.m.). Second, there was a non-monotonic association between baseline pupil diameter and decision bias in mice (McGinley, David, et al., 2015), in contrast to the monotonic pattern we observed here for phasic arousal in the same dataset (Fig. 1F). Third, while the pupil responses were negatively related to baseline pupil diameter in the basic yes/no ($r = -0.159$, ± 0.017 s.e.m.), the yes/no rare condition ($r = -0.163$, ± 0.041 s.e.m.), the yes/no frequent condition ($r = -0.109$, ± 0.047 s.e.m.), and numbers ($r = -0.482$, ± 0.010 s.e.m.) data sets, there was either no or a weak systematic (linear or non-monotonic) association between baseline pupil diameter and decision bias (go/no-go: $p = 0.975$; yes/no: $p = 0.557$; yes/no rare: $p = 0.043$; yes/no frequent: $p = 0.556$; numbers: $p = 0.289$). Fourth, in the yes/no recognition task, there was again a negligible correlation between pupil response and preceding baseline diameter ($r = -0.010$, ± 0.010 s.e.m.). Thus, the behavioral correlates of pupil responses reported in this paper reflect genuine effects of phasic arousal, which are largely uncontaminated by the baseline arousal state.

Discussion

Arousal is traditionally thought to globally up-regulate the efficiency of information processing (e.g., the quality of evidence encoding or the efficiency of accumulation (Aston-Jones & Cohen, 2005; McGinley, Vinck, et al., 2015)). However, recent work indicates that phasic arousal signals might have more specific effects, such as reducing the impact of prior expectations and biases on decision formation (de Gee et al., 2017, 2014; Krishnamurthy et al., 2017; Nassar et al., 2012; Urai et al., 2017). We here established a principle of the function of phasic arousal in decision-making, which generalizes across species (humans and mice) and behavioral tasks (from perceptual to memory-based and numerical decisions): suppressing maladaptive biases in the accumulation of evidence leading up to a choice.

We identified this principle in 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 accumulation of the fluctuating sensory input (evidence). We furthermore showed that phasic arousal flexibly reduces accumulation biases, irrespective of sign, in the presence of different stimulus statistics. Finally, we established the pupil-linked suppression of evidence accumulation bias also for memory-based decision, and for higher-level form of human bias widely known in behavioral economics: risk-seeking. We conclude that the ongoing deliberation culminating in a choice (Shadlen & Kiani, 2013) is shaped by transient boosts in the global arousal state of the brain, in a stereotyped fashion: suppression of evidence accumulation bias.

We here used pupil responses as a peripheral readout of changes in cortical arousal state (Larsen & Waters, 2018; McGinley, Vinck, et al., 2015). Indeed, recent work has shown that pupil diameter closely tracks several established measures of cortical arousal state (Larsen & Waters, 2018; McGinley, Vinck, et al., 2015). Changes in pupil diameter are also associated with locus coeruleus (LC) responses in humans (de Gee et al., 2017; Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014), monkeys (Joshi, Li, Kalwani, & Gold, 2016; Varazzani, San-Galli, Gilardeau, & Bouret, 2015), and mice (Breton-Provencher & Sur, 2019; Liu, Rodenkirch, Moskowitz, Schriver, & Wang, 2017; Reimer et al., 2016). But some of these studies also found unique contributions to pupil size in other subcortical regions like the superior and inferior colliculi, the cholinergic basal forebrain and dopaminergic midbrain (de Gee et al., 2017; Joshi et al., 2016; Reimer et al., 2016). Thus, we remain agnostic about the exact neuroanatomical source(s) of the reported effects of phasic arousal.

We chose the drift diffusion model to capture the behavioral data from 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 choices and shape of RT distributions, (ii) has been shown to successfully account for behavioral data from a wide range of decision-making tasks, including go/no-go (Ratcliff et al., 2016;

Ratcliff & McKoon, 2008), and (iii) is, under certain parameter regimens, equivalent to a reduction of biophysically detailed neural circuit models of decision-making (Bogacz et al., 2006; Wong & Wang, 2006). The drift diffusion 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, which is supported by neurophysiological data from yes/no tasks (Deco, Pérez-Sanagustín, de Lafuente, & Romo, 2007; Donner, Siegel, Fries, & Engel, 2009). Third, in the go/no-go task, that subjects set an implicit boundary for no-choices (Ratcliff et al., 2016). The quality of our model fits suggest that the model successfully accounted for the measured behavior, lending support to the conclusions drawn from the parameter estimates.

The behavioral models for the perceptual and value-based task included a somewhat different mechanism to account for the pupil-predicted bias suppression. First, the drift diffusion model accounted for the suppression of perceptual and memory choice by assuming an additive mechanism: arousal shaped the evidence-independent constant (towards "yes" or "no") that was added to the mean evidence (i.e., drift). Second, the selective integration model accounted for the reduction of a pro-variance (risk-seeking) bias by assuming a multiplicative mechanism: arousal shaped the multiplicative gain (weighting) of the momentary evidence. Could a single additive or multiplicative mechanism capture the bias shifts in both tasks? We think this is unlikely, as the overt bias was different in nature: the perceptual and memory choice biases were an overall tendency to responding "yes" more often than "no" (or the other way around) regardless of the objective evidence. By contrast, the pro-variance bias did not map onto the choice boundaries (for "left" and "right") but was an overall tendency to choose the more volatile sequence of values, and thus depends on the interaction with the evidence. Although the suppression of perceptual and memory choice biases was well accounted for by an additive effect, we cannot, at present, rule out a multiplicative effect on momentary evidence accumulation. From the analyses presented here, we can conclude that phasic arousal affects the evidence accumulation process, resulting in a monotonic suppression of a wide range of biases.

The monotonic effects of phasic arousal on decision biases that we report here contrast with recent observations that tonic (pre-stimulus) arousal levels have a non-monotonic (inverted U) effects on behavior (perceptual sensitivity and bias) and neural activity (the signal-to-noise ratio of thalamic and cortical sensory responses; (Gelbard-Sagiv, Magidov, Sharon, Hendler, & Nir, 2018; McGinley, David, et al., 2015). Importantly, our current work enables a direct comparison of these functional correlates of tonic and phasic arousal within the exact same data set in mice. A previous report on that data set showed that the mice's behavioral performance was most rapid, accurate, and the least biased at intermediate arousal (medium baseline pupil size) (McGinley, David, et al., 2015). 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 more 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).

Recent findings indicate that intrinsic behavioral variability is increased during sustained ("tonic") elevation of NA levels, in line with the "adaptive gain theory" (Aston-Jones & Cohen, 2005). First, optical stimulation of LC inputs to anterior cingulate cortex caused rats to abandon strategic counter prediction in favor of stochastic choice in a competitive game (Tervo et al., 2014). Second, chemogenetic stimulation of the LC in rats performing a patch leaving task increased decision noise and subsequent exploration (Kane et al., 2017). Third, pharmacologically reducing central noradrenaline levels in monkeys performing an operant effort exertion task parametrically increased choice consistency (Jahn et al., 2018). Finally, pharmacologically increasing central tonic noradrenaline levels in human subjects boosted the rate of alternations in a bistable visual input and long-range correlations in brain activity (Pfeffer et al., 2018). Here, we tested for the effect of phasic arousal on a range of behavioral parameters, including decision noise. In the drift diffusion model, increased decision noise would manifest as a decrease of the mean drift rate, which scales inversely with (within-trial) decision noise. We found no such effect that was consistent across data sets. This is another indication, together with the baseline pupil effects reported by (McGinley, David, et al., 2015), that the effects of phasic and tonic neuromodulation are distinct.

One influential account holds that phasic LC responses during decision-making are triggered by the threshold crossing in some circuit accumulating evidence, and that the resulting noradrenaline release then facilitates the translation of the choice into a motor act (Aston-Jones & Cohen, 2005). Within the drift diffusion model, this predicts a reduction in non-decision time and no effect on evidence accumulation. In contrast to this prediction, we found that in all our datasets that phasic arousal affected evidence accumulation (suppressing biased therein), but not non-decision time. Our approach does not enable us to rule out an effect of phasic arousal on movement execution (i.e., kinematics). Yet, our results clearly establish an important role of phasic arousal in evidence accumulation, ruling out any *purely* post-decisional account. This implies that phasic LC responses driving pupil dilation are already recruited during evidence accumulation, or that the effect of pupil-linked arousal on evidence accumulation are governed by systems other than the LC. Future experiments characterizing phasic activity in the LC or other brainstem nuclei involved in arousal during protracted evidence accumulation tasks could shed light on this issue.

It is tempting to speculate that task-evoked neuromodulatory responses and cortical decision circuits 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; Hasselmo, 2006; Hsieh, Cruikshank, & Metherate, 2000; Kimura, Fukuda, & Tsumoto, 1999; Kobayashi et al., 2000). Sensory cortical regions encode likelihood signals and sent these (bottom-up) to association cortex; participants' prior beliefs (about for example target probability) are sent back (top-down) to the lower levels of the hierarchy (Beck, Ma, Pitkow, Latham, & Pouget, 2012; Pouget, Beck, Ma, & Latham, 2013). Neuromodulators might reduce the weight of this prior in the inference process (Friston, 2010; Moran et al., 2013), thereby reducing choice biases. Another possibility is neuromodulator release might scale with uncertainty about the incoming sensory data (Friston, 2010; Moran et al., 2013). Such a process could be implemented by top-down control of the cortical systems decision-making over neuromodulatory brainstem centers. This line of reasoning is consistent with anatomical connectivity (Aston-Jones & Cohen, 2005; Sara, 2009). Finally, a related conceptual model that has been proposed for phasic LC responses is that cortical regions driving the LC (e.g. ACC) continuously compute the ratio of the posterior probability of the state of the world, divided by its (estimated) prior probability (Dayan & Yu, 2006). LC is then activated when neural activity ramps towards the non-default choice (against ones' bias). The resulting LC activity might reset its cortical target circuits (Bouret & Sara, 2005) and override the default state (Dayan & 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 & Delgado, 2017). For example, acute stress reduces risk-seeking when making decisions involving financial gains (Porcelli & Delgado, 2009), it increases overexploitation in sequential foraging decisions (Lenow, Constantino, Daw, & Phelps, 2017), and it impairs "model-based" goal-directed choice behavior (Otto, Raio, Chiang, Phelps, & Daw, 2013; Schwabe, Hoffken, Tegenthoff, & Wolf, 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.

If, as we show, phasic arousal monotonically optimizes evidence accumulation, why is not always engaged as strongly as possible? There are several possible reasons. First, phasic LC responses depend in a non-monotonic fashion on baseline LC activity (Aston-Jones & Cohen, 2005). The baseline activity of neuromodulatory brainstem centers, in turn, is shaped by the many inputs that they receive. Most

inputs inform about general bodily state, and the top-down inputs conveying the cognitive signals related to decision-making make up only a modest fraction of these inputs. Thus, even if the top-down signals driving phasic arousal were perfectly calibrated to continuously optimize task performance, the phasic arousal responses per se might not be. In other words, arousal systems might well act close to optimally in juggling all a plethora of tradeoffs between different tasks in real life; our optimality analysis only focuses on a small subset. Related, neuromodulatory baseline activity likely mediates shifts between rest, exploitation (performing the task in order to obtain the rewards) and exploration (looking for more rewarding alternatives) (Aston-Jones & Cohen, 2005; McGinley, Vinck, et al., 2015). Thus, through the same baseline dependence, phasic arousal might not always be perfectly calibrated to optimize task performance. Finally, phasic arousal might not always be engaged as strongly as possible because of its energetic costs: in the brain, neuromodulatory activation of G-protein coupled cascades are metabolically costly, and taxing on cells due to for example free radicals, and calcium load; in the rest of the body, LC-NA signaling triggers the release of glucose from energy stores.

Our study showcases the value of comparative experiments in humans and non-human species. One would expect the basic functions of arousal systems (e.g. the LC-NA system) to be analogous in humans and rodents. Yet, it has been unclear whether these systems are recruited by the same computational variables entailed in decision-making. Computational variables like decision uncertainty or surprise are encoded in prefrontal cortical regions (e.g. anterior cingulate or orbitofrontal cortex (Kepecs, Uchida, Zariwala, & Mainen, 2008; Ma & Jazayeri, 2014; Pouget, Drugowitsch, & Kepecs, 2016) and conveyed to brainstem arousal systems via top-down projections (Aston-Jones & Cohen, 2005; Breton-Provencher & Sur, 2019). Both the cortical representations of computational variables and top-down projections to brainstem may differ between species. More importantly, it has been unknown whether key components of the decision formation process, in particular evidence accumulation, would be affected by arousal signals in the same way between species. Only recently has it been established that rodents (rats) and humans accumulate perceptual evidence in an analogous fashion (Brunton et al., 2013). Here, we established that the shaping of evidence accumulation by phasic arousal is also governed by a conserved principle.

MATERIALS AND METHODS

Subjects

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, David, et al., 2015). 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 task), the ethics committee of Baylor College of Medicine (yes/no task with biased

signal probabilities), the ethics committee of the University of Hamburg (recognition task) 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 (yes/no task with biased signal probabilities, recognition task and value-based choice task) for their participation. We analyzed three previously unpublished human data sets, and re-analyzed a previously published mice data set (McGinley, David, et al., 2015) and two human data sets (Bergt, Urai, Donner, & Schwabe, 2018; de Gee et al., 2017). Bergt et al. (2018) have analyzed pupil responses only during the encoding phase of the recognition memory experiment – we here present the first analyses of pupil responses during the recognition phase.

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. Fifteen human subjects (8 females; age range, 20–28 y) performed the yes/no task with biased signal probabilities. Fifty-four human subjects (27 females; age range, 18–35 y) performed a picture recognition task, of which two were excluded from the analyses due to eye-tracking failure. 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 tracking data quality and/or excessive blinking.

For the go/no-go task, mice performed between five and seven sessions (described in (McGinley, David, et al., 2015)), 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 yes/no task with biased signal probabilities, human subjects performed 8 blocks of 120 trials each (distributed over two measurement sessions), yielding a total of 960 trials per participant. For the picture recognition task, human subjects performed 300 trials. 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, David, et al., 2015), or a pure sine wave (2 KHz) superimposed onto one of the noise stimuli. 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 behavioral set-up as in (de Gee et al., 2014).

Mice learned to respond during the signal-plus-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 μL 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 loudness 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 loudness exhibited a robust effect on mean accuracy, with highest accuracy for the loudest signal level: $F(5,20) = 23.95$, $p < 0.001$) and $F(4,76) = 340.9$, $p < 0.001$), for mouse and human subjects respectively. Human hit-rates were almost at ceiling level for the loudest signal: 94.7% (±0.69% s.e.m.). Because so few errors are not enough to sufficiently constrain the drift diffusion model, we merged the two conditions with the loudest signals.

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

Each trial consisted of two consecutive intervals (Fig. 3A): (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, David, et al., 2015), or a pure sine wave (2 KHz) superimposed onto the noise. In the first experiment, the signal was presented on 50% of trials (Fig. 3A). Auditory stimuli were presented at the same intensity of 65dB using the same over-ear headphone as in the go/no-go task. In the second experiment, in order to experimentally manipulate perceptual choice bias, the signal was presented on either 30% of 70% of trials. Auditory stimuli were presented at approximately the same loudness (65dB) using a Sennheiser HD 660 S over-ear headphone, suppressing ambient noise.

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. In the second experiment, subjects were explicitly informed about signal probability. The order of signal probability (e.g., first 480 trials –> 30%; last 480 trials –> 70%) was counterbalanced across subjects.

Throughout the experiment, the target signal loudness was fixed at a level that yielded about 75% correct choices in the 50% signal probability condition. Each participant's individual loudness 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-plus-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). In the first experiment, the resulting threshold loudness produced a mean accuracy of 74.14% correct (±0.75 % s.e.m.). In the second experiment, the resulting threshold loudness produced a mean accuracy of 84.40% correct (±1.75% s.e.m.) and 83.37% correct (±1.36% s.e.m.) in the P(Signal)=0.3 and P(Signal)=0.7 conditions, respectively. This increased accuracy was expected given the subjects' ability to incorporate prior knowledge about signal probability into their decision-making.

Picture yes/no (forced-choice) recognition

The full experiment consisted of a picture and word encoding task, and a 24 hours-delayed free recall and recognition tests (Fig. 5A) previously described in (Bergt et al., 2018). Here we did not analyze data from the word recognition task because of a modality mismatch: auditory during encoding, visual during recognition. During encoding, 75 neutral and 75 negative greyscale pictures (modified to have the same average luminance) were randomly chosen from the picture pool (Bergt et al., 2018) and presented in randomized order for 3 seconds at the center of the screen, against a grey background that was equiluminant to the pictures. Subjects were instructed to memorize the pictures (intentional encoding) and to evaluate how emotional each picture was on a 4-point scale from 0 ("neutral") to 3 ("very negative"). During recognition, 24-hours post encoding, subjects saw all pictures that were presented on the first day and an equal number of novel neutral and negative items in randomized order. Subjects were instructed to indicate for each item whether it had been presented the previous day ("yes – old") or not ("no – new"). For items that were identified as "old", participants were further asked to rate on a scale from 1 ("not certain") to 4 ("very certain") how confident they were that the item was indeed "old".

Value-based choice task

Each trial consisted of four consecutive intervals (Fig. 6A): (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 $\mu L = \mu H - d$. 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 integration 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.

Pupil data acquisition

The mouse pupil data acquisition is described elsewhere (McGinley, David, et al., 2015). 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 time derivative of the pupil size, by subtracting the size from adjacent frames, and smoothened the resulting time series with a sliding boxcar window (width, 50 ms).

Quantification of task-evoked pupil responses

The auditory yes/no tasks and the yes/no recognition task were analogous in structure to the tasks from our previous pupillometry and decision-making studies (de Gee et al., 2017, 2014). We here computed task-evoked pupil responses time-locked to the behavioral report (button press). 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. Specifically, we computed 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 Figs. S3B, S4A, S5A). The resulting pupil bins were associated with different overall pupil response amplitudes across the whole duration of the trial (Figs. S3C, S4B, S5B).

The go/no-go and value-based choice task entailed several deviations from the above task structure that required different quantifications of task-evoked pupil responses. The go/no-task had, 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 to 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 approaches 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). For the mice, this window ranged from 40–190 after sound onset; for humans, this window ranged from 240–460 ms after sound onset. Second, we excluded decision intervals

with a motor response before the end of this window plus a 50 ms buffer (cutoff: 240 ms for mice, 510 ms for humans; Fig. S1E,K). In both species, the resulting pupil derivate defined bins were associated with different overall pupil response amplitudes across the whole duration of the trial (Fig. S1G,M).

In the value-based choice task, 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. We choose this window of interest for three reasons. First, to quantify the amplitude of phasic arousal across the full interval of evidence accumulation, which was substantially longer than in the go/no-go and yes/no tasks (4.0–6.4 s vs. ~1 s). Second, the first 1.5 s after stimulus onset were excluded because here we observed initial constriction of the pupil below pre-stimulus baseline level, likely elicited by the high-contrast numbers elicited (Fig. S6A). Third, 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 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 different overall pupil response amplitudes across the whole duration of the trial (Fig. S6B).

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 yes/no task with biased environments and the value-based task, because subjects performed substantially fewer trials (see *Subjects*). We used two bins for the recognition task, so that we could perform the individual difference analysis reported in Fig. 5. In the recognition task, we ensured that each pupil bin contained an equal number of neutral and emotional stimuli. In all cases, the results are qualitatively the same when using five equally populated bins of task-evoked pupil response amplitudes.

Analysis and modeling of choice behavior

In the go/no-go task, each stimulus in a given trial (i.e., sequence discrete signal-plus-noise or noise-only sounds) was interpreted 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. S1K).

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

The signal detection metrics sensitivity (d') and criterion (c) (Green & Swets, 1966) were computed separately for each of the bins of pupil response size. We estimated d' as the difference between z-scores of hit-rates and false-alarm rates. We estimated criterion by averaging the z-scores of hit-rates and false-alarm rates and multiplying the result by -1.

In the go/no-go task, subjects could set only one decision criterion (or bias set point) against which to compare sensory evidence, because loudness was drawn pseudo-randomly on each trial. Therefore, using signal detection theory, per pupil bin we computed an overall perceptual choice bias across loudness as follows. We computed one false alarm rate (based on the noise sounds) and multiple hit-rates (one per loudness). Based on these we modelled one overall noise distribution (normally distributed with mean=0, sigma=1), and one "composite" signal distribution (Fig. S1A), which was computed as the average across a number of signal distributions separately modelled for each loudness (each normally distributed with mean=empirical d' for that loudness, and sigma=1). Thus, the standard signal detection theory assumptions were applied for each stimulus.

We defined the "zero-bias point" *(Z)* as the value for which the noise and composite signal distributions crossed:

Eq.1:

$$S(Z) - N(Z) = 0$$

where *S* and *N* are the composite signal and noise distributions, respectively.

The subject's empirical "choice point" *(C)* was computed as:

Eq.2:

$$C = (0.5 \times d') + c$$

where *d'* and *c* are a subject's signal detection theoretic sensitivity and criterion for a given loudness. Note that, as *C* is a constant when d' and criterion are computed for each loudness based on the same false alarm rate, it does not matter which loudness is used to compute the empirical choice point.

Finally, the overall bias measure was then taken as the distance between the subject's choice point and the zero-bias point:

Eq.3:

$$\text{Overall bias} = C - Z.$$

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

For both the go/no-go data sets, we calculated one group-average false alarm rate, and one group-average hit-rate per loudness. As above, we computed d' separately per loudness as the difference between z-scores of hit- and false-alarm rates, while using the same false alarm rate for each loudness. We then generated one noise distribution (normally distributed with mean=0, sigma=1) and separate signal distributions for each loudness (normally distributed with mean=empirical group-average d' for that loudness). The noise distribution was three times larger than the signal-plus-noise distribution because subjects encountered more noise sounds (follows from the probabilities in Fig. S1A). We then simulated 1 million trials for a range of choice biases (SDT criterion). Criterion ranged from -3 to 3 in steps of 0.01. On each trial, signal position (#2-7 in the sequence) and loudness were drawn randomly as in the actual task. On every sound interval, depending on the randomly selected stimulus, the agent's internal decision variable (DV) was randomly drawn from the noise or from one of the signal+noise distributions. 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. For every criterion value we computed reward rate as the number of reward divided by the total time to complete the one million trials, and we recomputed our overall bias measure based on the observed false alarm rate and hit-rates (see *Signal-detection theoretic modeling (go/no-go and yes/no tasks).* Optimality was defined as the overall bias value that maximized reward rate (# rewards / total time) (Fig. S1D,J) .

Drift diffusion modeling

Data from all tasks except the value-based choice task were exclusively fit with the drift diffusion model, which well captured all features of behavior we assessed. The value-based choice task was also

fit with more complex evidence accumulation models described below, because the standard drift diffusion model failed to capture the specific set of behavioral signatures of biased evidence accumulation in this task (Fig. 6F).

We used the HDDM 0.6.1 package (Wiecki, Sofer, & Frank, 2013) to fit behavioral data from the yes/no and go/no-go tasks. In all datasets, we allowed the following parameters to vary with pupil response-bins: (i) the separation between both bounds (i.e. response caution); (ii) the mean drift rate across trials; (iii) drift bias (an evidence independent constant added to the drift); (iv) the non-decision time (sum of the latencies for sensory encoding and motor execution of the choice). In the datasets using yes/no protocols, we additionally allowed starting point to vary with pupil response bin. In the go/no-go datasets, we allowed non-decision time, drift rate, and drift bias to vary with signal strength (i.e., loudness). The specifics of the fitting procedures for the yes/no and go/no-go protocols are described below.

To verify that best-fitting models indeed accounted for the pupil response-dependent changes in behavior, we generated a simulated data set using the fitted drift diffusion model parameters. Separately per subject, we simulated 100000 trials for each pupil bin (and, for the go/no-go data, for each loudness), while ensuring that the fraction of signal+noise vs. noise trials matched that of the empirical data; we then computed RT, and signal detection d' and overall bias (for the go/no-go data sets) or criterion (for the rest) 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 bias) would appear as random trial-to-trial variability in the accumulation process (drift rate variability) (Fig. 2E,H). For simplicity, we then pooled across loudness and simulated 100000 trials from two conditions that differed according to the fitted drift bias (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 bias model as described above to the simulated data, and another version of the model in which we fixed drift bias across the two conditions.

Yes-no task. We fitted all yes/no datasets using Markov-chain Monte Carlo sampling as implemented in the HDDM toolbox (Wiecki et al., 2013). 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. Bayesian MCMC generates full posterior distributions over parameter estimates, quantifying not only the most likely parameter value but also the uncertainty associated with that estimate. The hierarchical nature of the model assumes that all observers in a dataset are drawn from a group, with specific group-level prior distributions that are informed by the literature. In practice, this results in more stable parameter estimates for individual subjects, who are

constrained by the group-level inference. The hierarchical nature of the model also minimizes risks to overfit the data (Katahira, 2016; Vandekerckhove, Tuerlinckx, & Lee, 2011; Wiecki et al., 2013). Together, this allowed us to simultaneously vary all main parameters with pupil bin: starting point, boundary separation, drift rate, drift bias and non-decision time. We fixed drift rate variability across the pupil-defined bins. We ran 3 separate Markov chains with 12500 samples each. Of those, 2500 were discarded as burn-in. Individual parameter estimates were then estimated from the posterior distributions across the resulting 10000 samples. All group-level chains were visually inspected to ensure convergence. Additionally, we computed the Gelman-Rubin $\widehat{R}$ statistic (which compares within-chain and between-chain variance) and checked that all group-level parameters had an $\widehat{R}$ between 0.99-1.01.

Go/no-go task. The above described hierarchical Bayesian fitting procedure was not used for the go/no-go tasks because a modified likelihood function was not yet successfully implemented in HDDM. Instead, we fitted the go/no-go data based on RT quantiles, using the so-called G square method (code contributed to the master HDDM repository on Github; https://github.com/hddm-devs/hddm/blob/master/hddm/examples/gonogo_demo.ipynb). The RT distributions for yes-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; a single bin containing the number of no-go-choices contributed to G square (Ratcliff et al., 2016). Starting point and drift rate variability were fitted but fixed across the pupil-defined bins. Additionally, drift rate, drift bias and non-decision time varied with loudness. The same noise only intervals were re-used when fitting the model to each loudness.

The absence of no-responses in the go/no-go protocol required fixing one of the two bias parameters (starting point or drift bias) as function of pupil response. We fixed starting point based on formal model comparison between a model with pupil-dependent variation of drift bias and starting point: BIC differences ranged from -279.5 to -137.9 (mean, -235.3: median, -246.6), and from -197.5 to -146.0 (mean, -164.0; median, -162.0) in favor of the model with fixed starting point, for mice and humans respectively. The same was true when ignoring loudness: delta BICs ranged from -38.5 to -25.9 (mean, -30.9; median, -29.7), and from -39.8 to -26.7 (mean, -30.9; median, -30.7), for mice and humans respectively.

Modeling behavior from the value-based choice task

In the value-based task we modeled choice behavior using discrete-time accumulator models, complying with the fact pairs of numerical were presented at discrete time points. In all models described below two accumulators ($Y_{A,B}$, one per choice alternative) integrate numerical information over time (t). The two 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.

We fitted the following four models: a diffusion model (which is the discrete-time analogue of the drift-diffusion model), a leaky accumulation model (LAM), a leaky competing accumulator model (LCAM) and a leaky selective accumulator model (LSAM).

In the diffusion model the accumulators evolve over time according to the following difference equations:

Eq.4:

$$\mathrm{Y}_A(t) = \mathrm{Y}_A(t-1) + \mathrm{S}_A(t) + \xi \cdot \zeta_A(t)$$

$$\mathrm{Y}_B(t) = \mathrm{Y}_B(t-1) + \mathrm{S}_B(t) + \xi \cdot \zeta_B(t)$$

In the above, $S_{A,B}(t)$ are the inputs to the two accumulators on a given time-step, ξ is the standard deviation of the noise, and $\zeta_{A,B}(t)$ is the standard Gaussian samples (independent from each other and across time). The only free parameter of the model is the standard deviation of the noise ξ.

In the leaky accumulation model (LAM) the accumulators evolve according to:

Eq.5:

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

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

Relative to the diffusion model, the LAM has one extra parameter, λ which is the leak of the accumulation process. For λ=1 the LAM reduces to the diffusion model, whereas for $\lambda > 0$ ($\lambda < 0$) the model assigns larger weights to late (early) information, yielding thus a "recency" ("primacy") effect.

In addition to leak, the leaky competing accumulator model (LCAM) has a third parameter, β, which implements lateral inhibition between the two accumulators. Furthermore, in the LCAM the state of the accumulators cannot take negative values:

Eq.6:

$$\mathrm{Y}_A(t) = \max\left(0, (1-\lambda) \cdot \mathrm{Y}_A(t-1) - \beta \cdot \mathrm{Y}_B(t-1) + \mathrm{S}_A(t) + \xi \cdot \zeta_A(t)\right)$$

$$\mathrm{Y}_B(t) = \max\left(0, (1-\lambda) \cdot \mathrm{Y}_B(t-1) - \beta \cdot \mathrm{Y}_A(t-1) + \mathrm{S}_B(t) + \xi \cdot \zeta_B(t)\right)$$

Finally, the leaky selective accumulator model (LSAM) has also three free parameters:

Eq.7a:

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

$$Y_B(t) = (1-\lambda)\cdot Y_B(t-1) + I_B(t) + \xi \cdot \xi_B(t)$$

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.7b:

$$I_A(t) = \theta(S_A(t), S_B(t)) \cdot S_A(t)$$

$$I_B(t) = \theta(S_B(t), S_A(t)) \cdot S_B(t)$$

with function θ as follows:

Eq.7c:

$$\theta(x, y) = \frac{1}{1 + e^{-w(x-y)}}$$

This logistic function returns a value of 0.5 when the inputs are equal ($x = y$) and a value larger than 0.5 when $x > y$, which gets larger the larger the difference between x *and* y is. Parameter w is the selective gain parameter that controls the slope of the function. The larger the selective gain is the stronger the selective modulation will be, while for $w = 0$ the LSAM reduces to the LAM (Glickman, Tsetsos, & Usher, 2018).

We fitted the four models above using a maximum likelihood approach together with Bayesian BIC comparisons (Tsetsos et al., 2016). The actual stochastic sequences that participants encountered in the experimental trials were fed as input to the models. Predictions from the diffusion, LAM and LSAM were derived numerically while the LCAM model was simulated (1000 times per trial). Finding the maximum likelihood parameters for each model 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. For LSAM, which was the best-fitting model according to BIC comparisons, the model was also fitted to data from the pro-variance trials only and separately for each pupil bin.

Finally, in addition to quantitative model comparison, we pitted the models against two characteristic behavioral signatures obtained in the task (Fig. 6). First, the pro-variance bias (Tsetsos et al., 2012, 2016), which was defined as the as the fraction of high-variance choices across both trial types (sequence *A* when $\sigma A > \sigma B$; or sequence *B* when $\sigma A < \sigma B$. Second, the recency bias which was defined

by (i) estimating via logistic regression the weight assigned to the evidence at each time-point, and (ii) fitting these logistic weights using an exponential function with the sign of the exponent determining the type of the temporal-order bias (primacy/ recency).

Determining optimal choice bias in the value-based choice task

To understand whether trends in behavior related to pupil improved or degraded overall performance, we used the LSAM and fitted the data to all trials, regardless of pupil response. For the per participant best-fitting noise (ξ) and leak (λ) parameters we estimated the selective gain parameter (w) that achieved maximum accuracy (note that for $\xi > 0$ the accuracy maximizing w will also be larger than 0, see (Tsetsos et al., 2016)). Using the optimal w level, we predicted the optimal pro-variance bias (i.e., the bias maximizing percentage correct) using the optimal selective gain and the best-fitting noise (ξ) and leak (λ) parameters 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, Mächler, Bolker, & Walker, 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 loudness. Our approach was analogous to sequential polynomial regression analysis (Draper & 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.8:

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

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

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

with β as regression coefficients, S as the loudness (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, Levy, Scheepers, & Tily, 2013). For data from the go/no-go task, the random effects were specified to accommodate loudness coefficient to vary with participant, and the intercept and TPR-

coefficients to vary with loudness 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, Brockhoff, & Christensen, 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, Anne Bergt for help with the data collection for the human memory recognition task, 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, Investigation, Formal analysis, Writing—original draft, Writing—review and editing; KT, Conceptualization, Investigation, Formal analysis, 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

Amaral, D., & Sinnamon, H. (1977). The locus coeruleus: neurobiology of a central noradrenergic nucleus. *Progress in Neurobiology*, *9*(3), 147–196. https://doi.org/10.1016/0301-0082(77)90016-8

Arnsten, A. F. T. (2015). Stress weakens prefrontal networks: molecular insults to higher cognition. *Nature Neuroscience*, *18*(10), 1376–1385. https://doi.org/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*(1), 403–450.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, *67*(1).

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*(1), 30–39.

Bergt, A., Urai, A. E., Donner, T. H., & Schwabe, L. (2018). Reading memory formation from the eyes. *European Journal of Neuroscience*, *47*(12), 1525–1533. https://doi.org/10.1111/ejn.13984

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

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

Bouret, S., & Sara, S. J. (2005). Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences*, *28*(11), 574–582.

Bowen, H. J., Spaniol, J., Patel, R., & Voss, A. (2016). A Diffusion Model Analysis of Decision Biases Affecting Delayed Recognition of Emotional Stimuli. *PLOS ONE*, *11*(1), e0146769. https://doi.org/10.1371/journal.pone.0146769

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

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

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science (New York, N.Y.)*, *340*(6128), 95–98.

Colizoli, O., de Gee, J. W., Urai, A. E., & Donner, T. H. (2018). Task-evoked pupil responses reflect internal belief states. *Scientific Reports*, *8*(1), 13702. https://doi.org/10.1038/s41598-018-31985-3

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network (Bristol, England)*, *17*(4), 335–350.

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

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(5), E618–25.

Deco, G., Pérez-Sanagustín, M., de Lafuente, V., & Romo, R. (2007). Perceptual detection as a dynamical bistability phenomenon: A neurocomputational correlate of sensation. *Proceedings of the National Academy of Sciences*, *104*(50), 20073–20077.

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Current Biology : CB*, *19*(18), 1581–1585.

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

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, *92*(6), 1398–1411. https://doi.org/10.1016/j.neuron.2016.11.005

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Froemke, R. C. (2015). Plasticity of Cortical Excitatory-Inhibitory Balance. *Annual Review of Neuroscience*, *38*(1), 195–219. https://doi.org/10.1146/annurev-neuro-071714-034002

Gelbard-Sagiv, H., Magidov, E., Sharon, H., Hendler, T., & Nir, Y. (2018). Noradrenaline Modulates Visual Perception and Late Visually Evoked Activity. *Current Biology*, *28*(14), 2239-2249.e6. https://doi.org/10.1016/j.cub.2018.05.051

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

Glickman, M., Tsetsos, K., & Usher, M. (2018). Attentional Selection Mediates Framing and Risk-Bias Effects. *Psychological Science*, *29*(12), 2010–2019. https://doi.org/10.1177/0956797618803643

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

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

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

Hasselmo, M. E. (2006). The Role of Acetylcholine in Learning and Memory. *Current Opinion in Neurobiology*, *16*(6), 710–715. https://doi.org/10.1016/j.conb.2006.09.002

Hsieh, C. Y., Cruikshank, S. J., & Metherate, R. (2000). Differential modulation of auditory thalamocortical and intracortical synaptic transmission by cholinergic agonist. *Brain Research*, *880*(1–2), 51–64.

Hupé, J.-M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, *9*(7), 10.

Jahn, C. I., Gilardeau, S., Varazzani, C., Blain, B., Sallet, J., Walton, M. E., & Bouret, S. (2018). Dual contributions of noradrenaline to behavioural flexibility and motivation. *Psychopharmacology*, *235*(9), 2687–2702. https://doi.org/10.1007/s00213-018-4963-z

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

Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.

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

Katahira, K. (2016). How hierarchical models improve point estimates of model parameters at the individual level. *Journal of Mathematical Psychology*, *73*, 37–58. https://doi.org/10.1016/j.jmp.2016.03.007

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–231.

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. *The European Journal of Neuroscience*, *11*(10), 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*(5), e0155574.

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. *The European Journal of Neuroscience*, *12*(1), 264–272.

Krishnamurthy, K., Nassar, M. R., Sarode, S., & Gold, J. I. (2017). Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nature Human Behaviour*, *1*, 0107.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13).

Lak, A., Nomoto, K., Keramati, M., Sakagami, M., & Kepecs, A. (2017). Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. *Current Biology : CB*, *27*(6), 821–832.

Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits*, *12*.

Lee, C. R., & Margolis, D. J. (2016). Pupil Dynamics Reflect Behavioral Choice and Learning in a Go/NoGo Tactile Decision-Making Task in Mice. *Frontiers in Behavioral Neuroscience*, *10*, 200.

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

Lenow, J. K., Constantino, S. M., Daw, N. D., & Phelps, E. A. (2017). Chronic and Acute Stress Promote Overexploitation in Serial Decision Making. *The Journal of Neuroscience*, *37*(23), 5681–5689. https://doi.org/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 Reports*, *20*(13), 3099–3112.

Ma, W. J., & Jazayeri, M. (2014). Neural coding of uncertainty and probability. *Annual Review of Neuroscience*, *37*, 205–220.

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

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

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. *The Journal of Neuroscience*, *33*(19), 8227–8236.

Murphy, P. R., Boonstra, E., & Nieuwenhuis, S. (2016). Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nature Communications*, *7*, 13526. https://doi.org/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. *Human Brain Mapping*, *35*(8), 4140–4154.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, *110*(52), 20941–20946. https://doi.org/10.1073/pnas.1312011110

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

Pfeffer, T., Avramiea, A.-E., Nolte, G., Engel, A. K., Linkenkaer-Hansen, K., & Donner, T. H. (2018). Catecholamines alter the intrinsic variability of cortical population activity and perception. *PLOS Biology*, *16*(2), e2003453. https://doi.org/10.1371/journal.pbio.2003453

Porcelli, A. J., & Delgado, M. R. (2009). Acute Stress Modulates Risk Taking in Financial Decision Making. *Psychological Science*, *20*(3), 278–283. https://doi.org/10.1111/j.1467-9280.2009.02288.x

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. https://doi.org/10.1016/j.cobeha.2016.11.015

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nature Neuroscience*, *16*(9), 1170–1178.

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2016). Modeling Individual Differences in the Go/No-Go Task With a Diffusion Model. *Decision*.

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

Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, *84*(2), 355–362.

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

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223.

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

Schwabe, L., Hoffken, O., Tegenthoff, M., & Wolf, O. T. (2011). Preventing the Stress-Induced Shift from Goal-Directed to Habit Action with a -Adrenergic Antagonist. *Journal of Neuroscience*, *31*(47), 17317–17325. https://doi.org/10.1523/JNEUROSCI.3304-11.2011

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

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

Shadlen, M. N., & Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. *Neuron*, *90*(5), 927–939. https://doi.org/10.1016/j.neuron.2016.04.036

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

Tervo, D. G. R., Proskurin, M., Manakov, M., Kabra, M., Vollmer, A., Branson, K., & Karpova, A. Y. (2014). Behavioral Variability through Stochastic Choice and Its Gating by Anterior Cingulate Cortex. *Cell*, *159*(1), 21–32. https://doi.org/10.1016/j.cell.2014.08.037

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), 9659–9664.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(11), 3102–3107.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)*, *185*(4157), 1124–1131.

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, *8*, 14637.

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

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44–62. https://doi.org/10.1037/a0021765

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

Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, *86*(3), 740–754.

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

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*, 14.

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *The Journal of Neuroscience*, *26*(4), 1314–1328.

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

SUPPLEMENTARY FIGURES

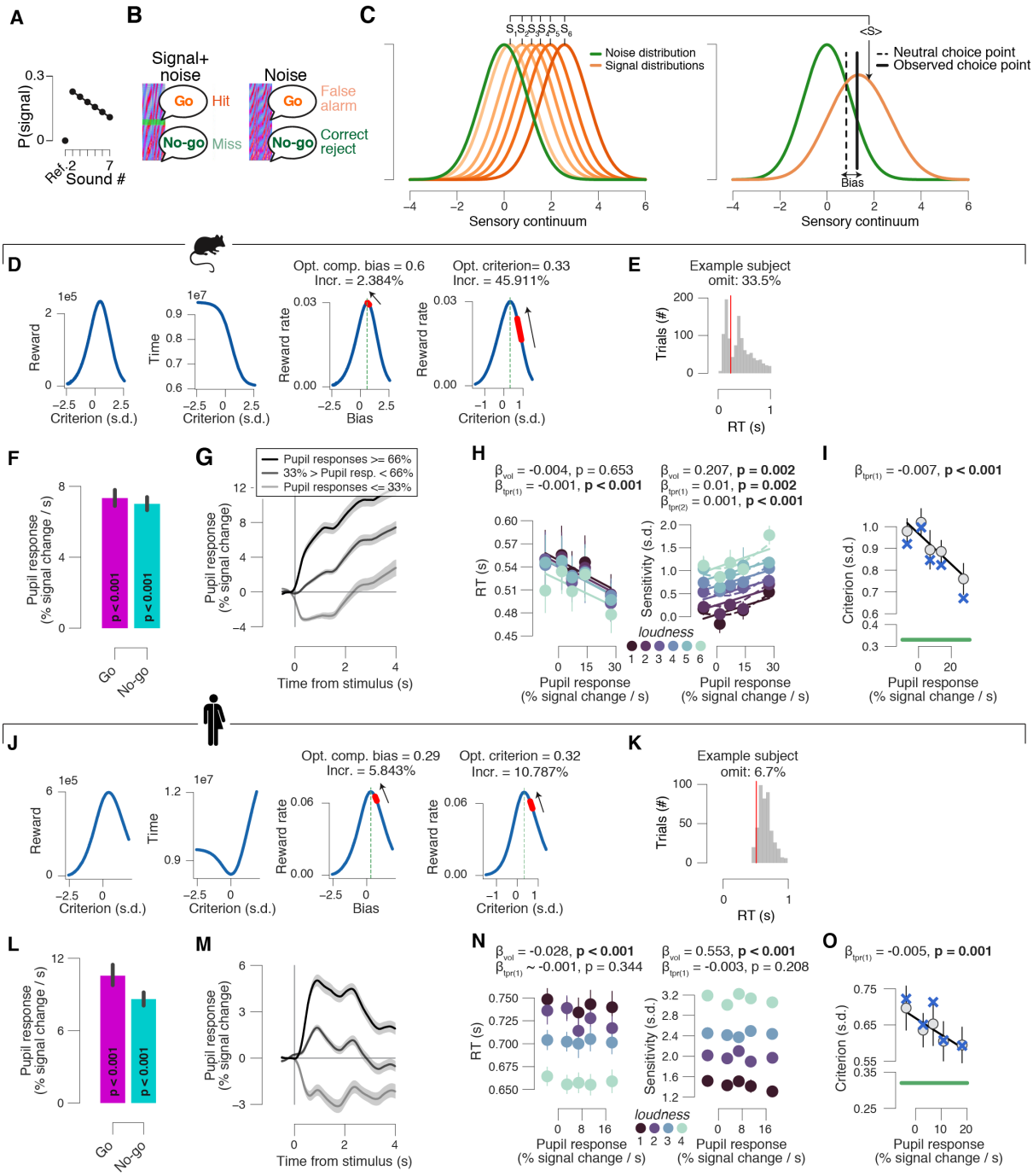

Figure S1. Quantifying pupil responses and behavior in mice and humans. (A) Probability of target signal across the sequence of 1–7 sounds. Hazard rate of signal occurrence was kept approximately flat. **(B)** The four combinations of stimulus category (signal+noise vs. noise) and behavioral choice (go vs. no-go) yielded the four standard signal detection theory categories. **(C)** Schematic of overall perceptual choice bias measure (Materials and Methods). Per pupil bin we modelled one overall noise distribution (green; normally distributed with mean=0, sigma=1), and one "composite" signal distribution. This

composite signal distribution was computed as the average across a number of signal distributions separately modelled for each loudness (orange; each normally distributed with mean=empirical d' for that loudness, sigma=1). We defined the "zero-bias point" *(Z)* as the value for which the noise and composite signal distributions cross. The subject's empirical "choice point" was computed based on the empirical d' and criterion for any difficulty level (Materials and Methods). The overall bias measure was then taken as the distance between the subject's choice point and the zero-bias point. (**D)** Results of simulation study of optimal bias in go/no-go task for mice (Materials and Methods). Optimality was defined as the level of overall choice bias (computed as in panel A; second from right) or average criterion (across loudness; right) that maximized reward rate (# rewards / total time). Red lines indicate range of group-average exhibited biased as a function of pupil response bin; arrows indicate direction from low to high pupil response bins. **(E)** RT distribution of example subject. Red line, 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). Range of omitted trials across all subjects: 29.9%–45.4% (mean, 35.3%; median, 33.5%). **(F)** Task-evoked pupil responses in mice sorted into go- and no-go choices (pooled across loudness). Stats, paired-samples t-test. **(G)** Overall pupil response time courses in mice for three pupil derivative defined bins (pooled across loudness). **(G)** Relationship between median RT (left), perceptual sensitivity (right; quantified by signal detection d') and pupil response in mice, separately for each loudness. 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. **(F)** As F, but for average criterion collapsed across loudness. 'X' markers are predictions from best fitting variant of drift diffusion model (Materials and Methods). **(J-O)**, as D-I, but for humans. Range of omitted trials in panel K: 0%–30.7% (mean, 7.0%; median, 4.4%). All panels: group average (N = 5; N = 20); error bars or shading, s.e.m.

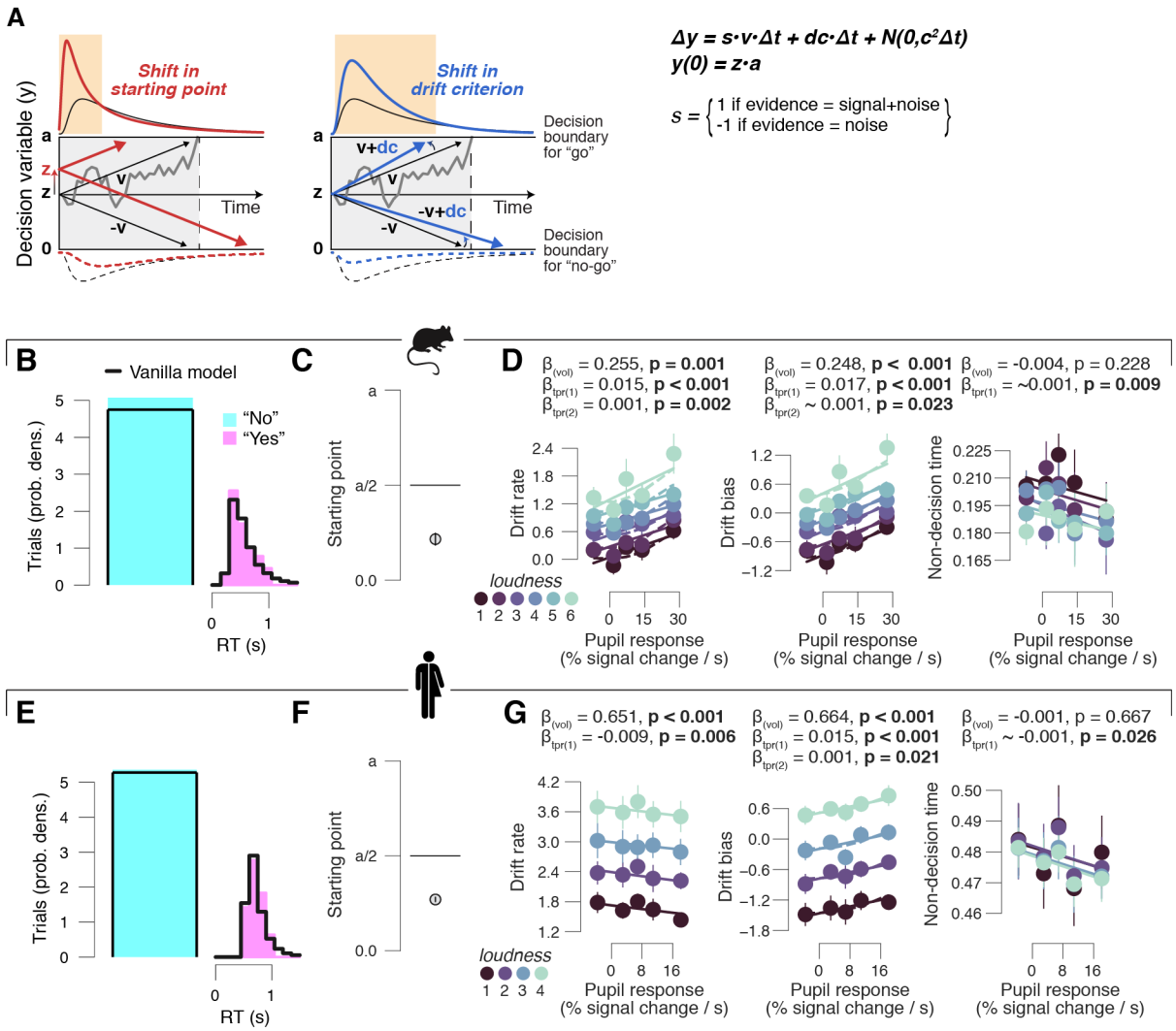

Figure S2. Pupil-dependent changes in computational model parameters during go/no-go task. **(A)** Schematic of drift diffusion model accounting for choices, and their associated RTs (for go-trials). Orange windows, RTs for which biased choices are expected under shifts in either "starting point" (*z*; left) or "drift bias" (*dc*; right). Solid (dashed) lines, (implicit) RT distributions. In the equation, *v* is the drift rate (estimated separately for each loudness). **(B)** Group average RT distributions, separately for yes- and no-choices. There were no RTs associated with no-choices (no-go); hence, a single bin containing the number of no-choices contributed to the model fit (Materials and Methods). Black lines, "vanilla model" fit (parameters boundary separation, drift rate, non-decision time, starting point and drift bias were fixed across loudness and pupil bins). **(C)** Starting point estimates of best fitting model (Materials and Methods) in mice expressed as a fraction of the boundary separation (*a*). **(D)** Relationship between drift rate estimates (left), drift bias estimates (right) of best fitting model (Materials and Methods) and pupil responses in mice, separately for each loudness. Linear fits are plotted wherever the first-order fit was superior to the constant fit. Quadratic fits were plotted (dashed

lines) wherever the second-order fit was superior to first-order fit. Stats, mixed linear modeling. **(E-G)** As B-D, but for humans. All panels: group average (N = 5; N = 20); error bars, s.e.m.

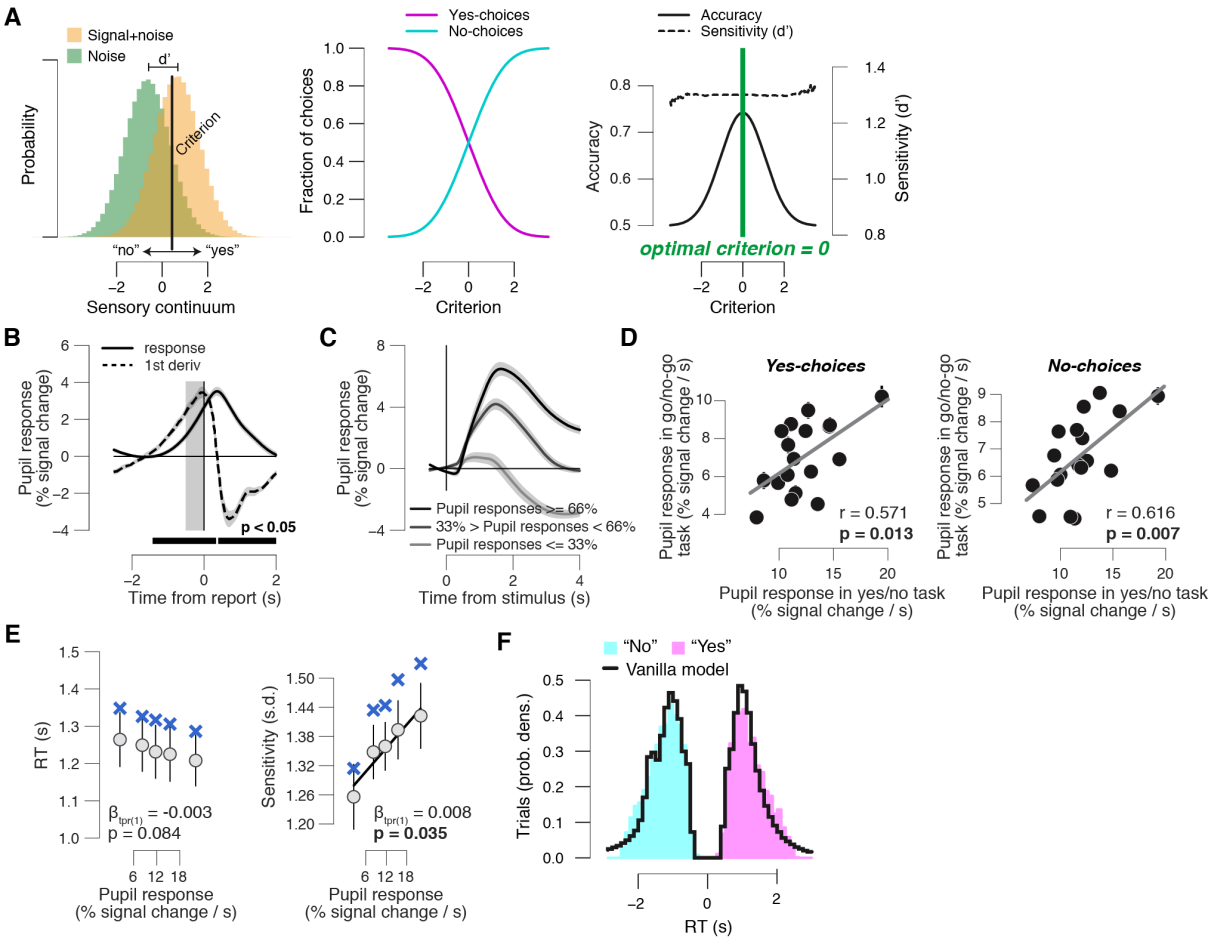

Figure S3. Optimality analysis of yes/no task. (A) Simulation study of optimal bias in yes/no (forced-choice) task (Materials and Methods). Optimality was defined as the level of choice bias (signal detection criterion) that maximized accuracy. Left: simulated internal sensory representations for signal+noise 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, loudness 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)** Overall pupil response time course for three pupil derivative defined bins. **(D)** 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. A leverage analysis verified that the reported correlations are not driven by outliers. **(E)** Relationship between RT (left), perceptual sensitivity (right) and pupil response. Linear fits are plotted wherever the first-order fit was superior to the constant fit. Quadratic fits were not superior to first-order fits. Stats, mixed-linear modeling. **(F)** Group average RT distributions, separately for yes- and no-choices. Black lines, "vanilla model" fit (parameters boundary separation, drift rate, non-decision time, starting point and drift bias were fixed across pupil bins.

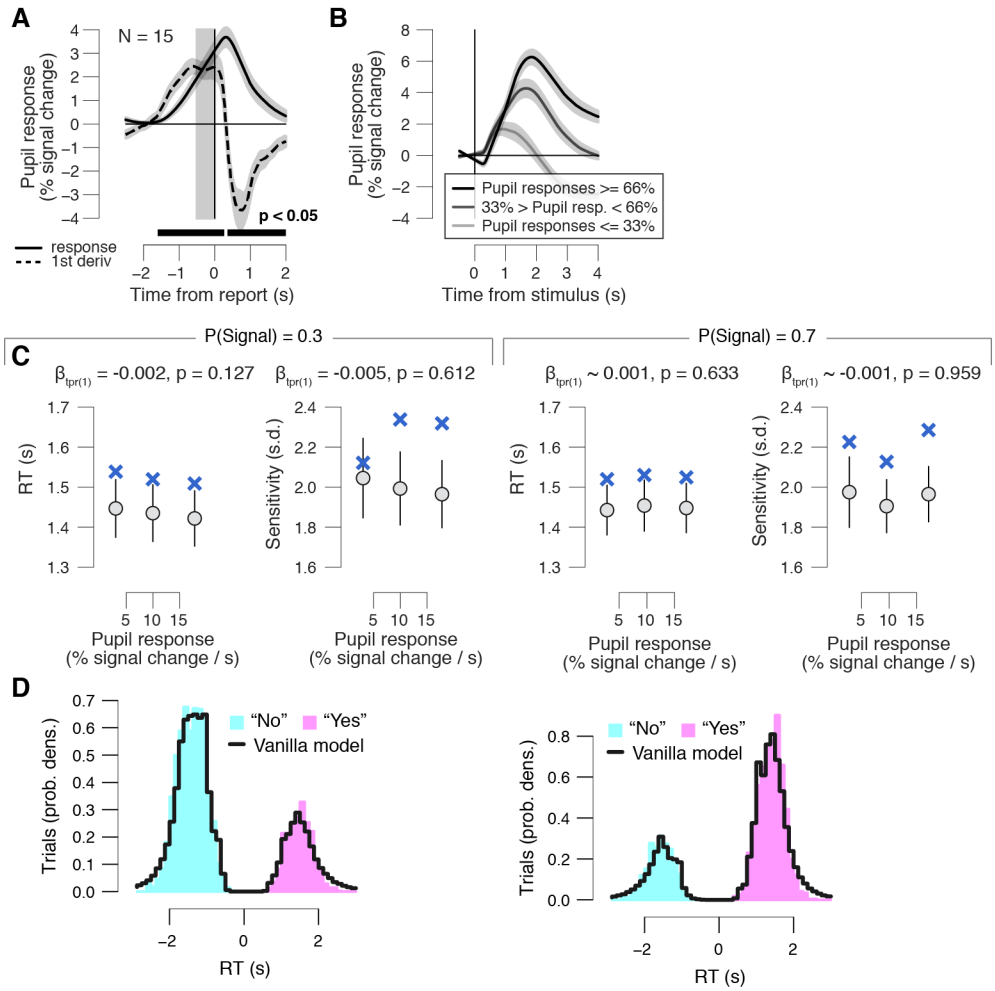

Figure S4. (A) Task-evoked pupil response (solid line) and response derivative (dashed line). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. Stats, paired-samples t-test. **(B)** Overall pupil response time course for three pupil derivative defined bins. **(C)** Relationship between RT, perceptual sensitivity (right) and pupil response in the rare condition (left) and frequent condition (right). Linear fits are plotted wherever the first-order fit was superior to the constant fit. Quadratic fits were not superior to first-order fits. Stats, mixed-linear modeling. **(D)** Group average RT distributions in the rare condition (left) and frequent condition (right). Black lines, "vanilla model" fit (parameters boundary separation, drift rate, non-decision time, starting point and drift bias were fixed across pupil bins).

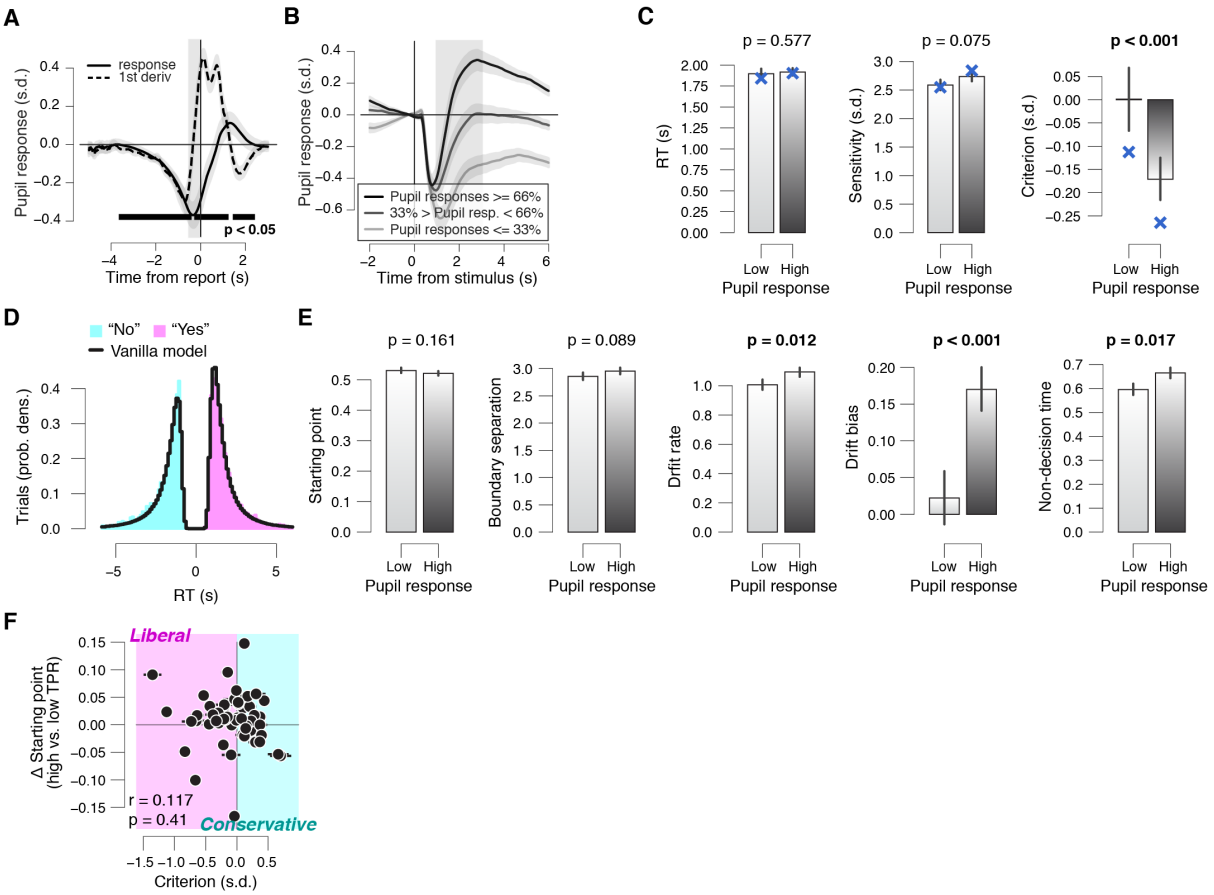

Figure S5. (A) Task-evoked pupil response (solid line) and response derivative (dashed line). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. Stats, paired-samples t-test. **(B)** Overall pupil response time course for three pupil derivative defined bins. Stats, paired-samples t-test. **(C)** RT (left), sensitivity (right), choice bias, (right) for low and high pupil response bins. **(D)** Group average RT distributions in the conservative condition, separately for yes- and no-choices. Black lines, "vanilla model" fit (parameters boundary separation, drift rate, non-decision time, starting point and drift bias were fixed across pupil bins). **(E)** As B, but for drift diffusion model parameters. **(F)** Individual pupil predicted shift in starting, plotted against individual's overall choice bias. Data points, individual subjects. Stats, Pearson's correlation. Error bars, 60% confidence intervals (bootstrap).

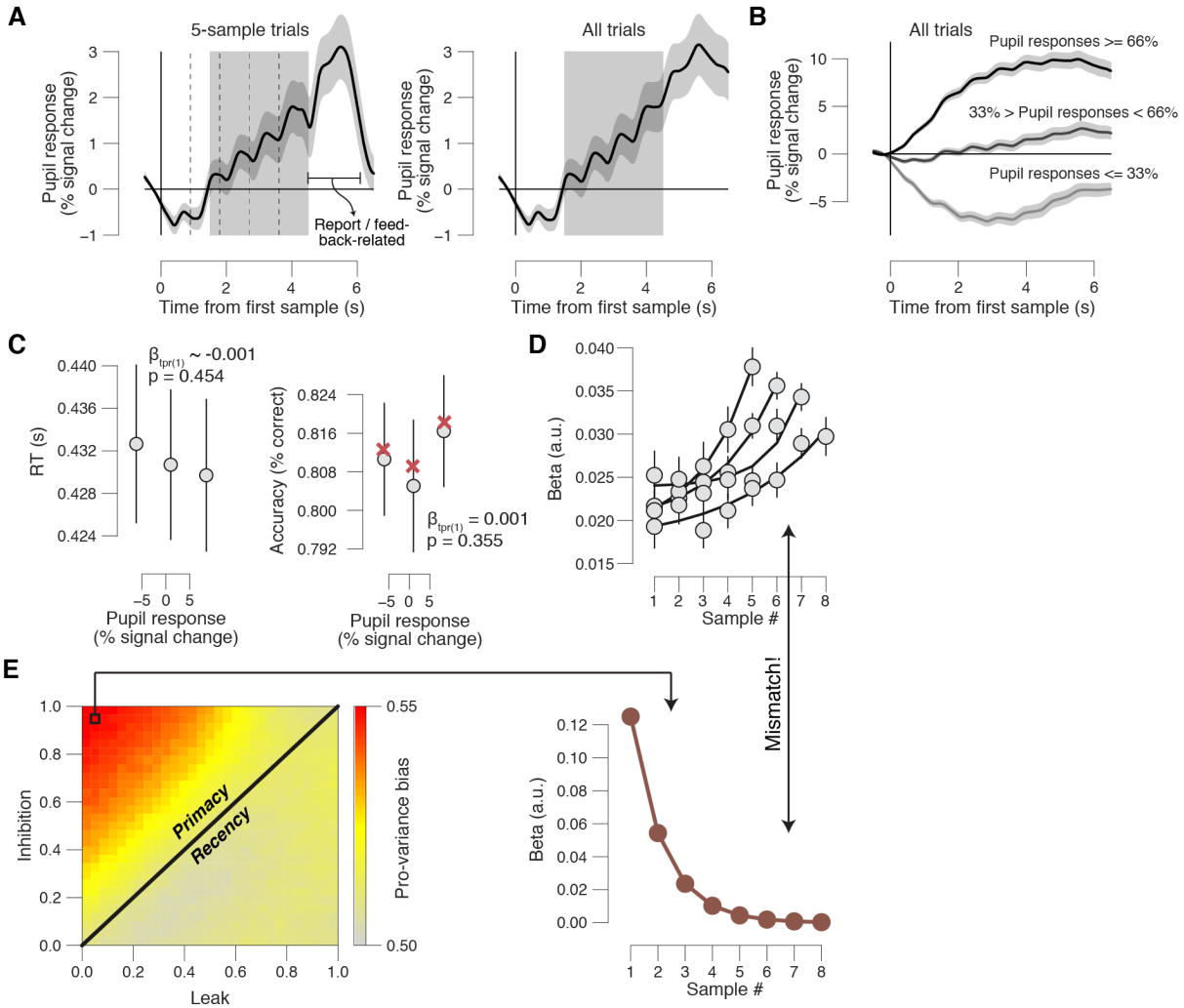

Figure S6. Pupil responses and pupil-dependent changes in behavior during value-based choice task. (A) Task-evoked pupil response time courses locked to the onset of the first pair of samples. Left: for the shortest trials (5 pairs of samples). Dashed vertical lines, onset of sample pairs 2-5. Right: for all trials (5–8 pairs of samples). Grey box, interval for computing single trial task-evoked pupil response measures (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 (Materials and Methods). Quadratic fits were not superior to first-order fits. Stats, mixed linear modelling. **(D)** Psychophysical kernels, indicating the effect of each pair of numbers on observers' choice, separately for each trial duration (5-8 samples). Lines, exponential fits to data. **(E)** Simulated pro-variance bias by the leaky competing accumulator model, for various levels of leak and mutual inhibition. The LCA can only produce pro-variance in an inhibition-dominant regime (mutual inhibition > leak). However, this produces a primacy bias (overweighing of early samples of evidence), which is

the opposite of what we found in the empirical data (see panel E). Therefore, the LCA does not provide a good fit to our behavioral data, while the selective integration model does (see main text). All panels: group average (N = 32); error bars or shading, s.e.m.