Universitätsklinikum Hamburg-Eppendorf

Zentrum für Experimentelle Medizin
Institut für Neurophysiologie und Pathophysiologie
Direktor: Prof. Dr. Andreas K. Engel

Neuromodulatory shaping of decision computations

Dissertation
zur Erlangung des Doktorgrades PhD
an der Medizinischen Fakultät der Universität Hamburg

vorgelegt von
Jan Willem de Gee
aus Amsterdam

Hamburg, 2019

Angenommen von der Medizinischen Fakultät am: 20.09.2019

Veröffentlicht mit Genehmigung der medizinischen Fakultät der Universität Hamburg

Prüfungsausschuss, der/die Vorsitzende:	Prof. Dr. Tobias Donner
Prüfungsausschuss, 2. Gutachter/in:	Prof. Dr. Philipp Sterzer
3. Gutachter:	Prof. Dr. Richard Ridderinkhof

The time is roughly fourteen billion years after the big bang. The place is a blue planet circling a medium-sized sun, located in the outskirts of the milky way. The Earth's rare combination of features – (i) not too far and not too close from the Sun for liquid water to exist, (ii) big enough to hold on to its atmosphere, (iii) a strong magnetic field to deflect harmful UV rays, and (iv) tectonic activity that helps feed the atmosphere and oceans with nutrients – enabled the emergence of life: organisms with the capacity to harness energy and materials from their environment to adapt to changing conditions and reproduce themselves. Life on Earth has evolved, and continues to do so, into ever more specialized plants and animals. A notable example are the homo sapiens – highly intelligent and curious creatures who are currently in the middle of a systematic attempt to understand the physical world they live in, including their own complexity. I hope this thesis may serve as a humble building block in this "grand scheme of things".

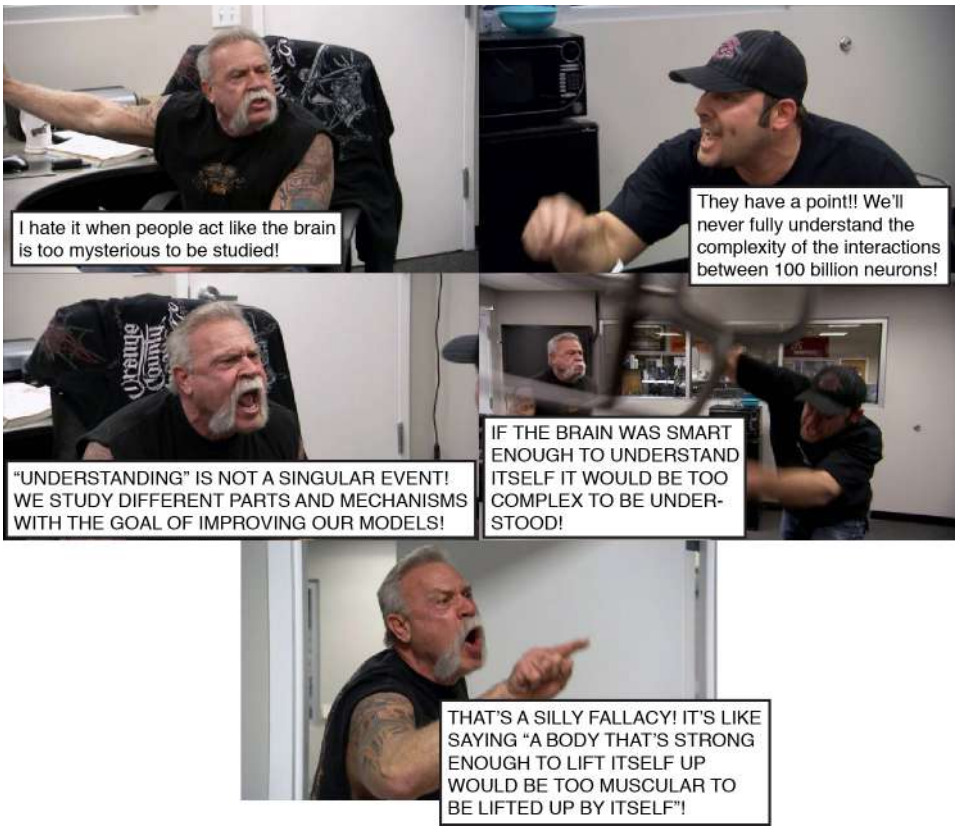


Figure 1: Images from American Chopper on Discovery Channel; meme by Dr. Jens Foell.

Contents

0 Synopsis

0.1 Introduction

Many decisions, ranging from the everyday to those with far-reaching consequences, have to be made in the face of uncertainty. Imagine walking down the street on your way to the nearest convenience store, when you suddenly stumble across a weekly open market. In the distance you see a fruit stall, and the items on display look attractive, although it is hard to be sure from afar. Do you decide to go and buy your fruits there, or do you continue your way to the grocery? Now imagine you are sitting in the oval office, it is 2002, and you are the president of the USA. Your intelligence officers just briefed you about the probability of weapons of mass destruction in Iraq. Do you decide to forcefully invade the region, or will you hold off for now?

How do we face this tremendous challenge of making sense of, and acting upon, the noisy world around us? One perhaps obvious approach is to base your decision not on one, but on at least a number of pieces of **"evidence"**. In the open market, one might combine (sum up) estimates of the fruit quality, the variety to choose from, and the extent to which the stall seems attractive to other people. In the oval office, one might combine satellite images with intelligence from informers on the ground, and also consult independent watchdog organizations and other world leaders. In both examples, the hope is that "evidence accumulation" will reduce uncertainty by (i) adding relevant information ("signal") and (ii) averaging out the error ("external noise") across samples of evidence.

Strikingly, when asked to make repeated decisions, we will often choose differently each time, even when we are given the same information to inform our choice [1–5]. This is called **"behavioral variability"**. Behavioral variability is typically attributed to spontaneous fluctuations of neural activity (**"internal noise"**) in the brain's decision-making machinery [1, 3]. This implies that the path a decision-making process takes in our brain is, to an extent, unpredictable. Indeed, George W. Bush might have decided against

invading Iraq if he had considered the same information on another occasion.

However, behavioral variability is perhaps not completely unpredictable: a substantial part of the variability in behavior might in fact be explained and predicted by tracking the momentary **"arousal state"** of the brain, and by tracking the **"history of choices"** a particular decision is embedded in.

The brain's global arousal state is largely governed by so-called neuromodulatory systems. Neuromodulatory systems originate in the brainstem and have widespread projections to the rest of the brain that tune the way their target circuits operate [6–10]. In the following, I will broadly refer to this neuromodulatory tuning of cortical circuit interactions as the "global arousal state". Neuromodulatory systems operate on different timescales. For example, activity of the locus coeruleus noradrenaline system is lowest during sleep, rises during wakefulness, and reaches much higher levels during situations of stress or danger; the locus coeruleus also activates briefly ("phasically") during cognitive acts like decision-making [6]. As a result, the arousal state of the brain constantly fluctuates [6, 11], and is slightly elevated every time we make a decision [6].

Choices are rarely made in isolation, but often depend on the sequence of previously made choices [12, 13]. This holds even for low-level perceptual choices [14–16]. This implies that the choice history can bias the brain's decision-making machinery. For example, the decision to invade Iraq might have biased George W. Bush's assessment of new information to confirm to his initial judgment – a phenomenon called confirmation bias [17].

Taken together, the brain's arousal state and the history of choices should, somehow, shape cortical decision-computations and subsequent choice behavior. The goal of this thesis is to determine how. To uncover how arousal signals and the history of choices tune decision-making processes in the brain, would mean that a significant fraction of the observed variance (variability) in choice behavior is not random but can be predicted.

One powerful starting point proved to be comparing decision-computations and subsequent choice behavior between relatively low and high arousal states. This approach relies heavily on participants making a large number of decisions that are under tight experimental and analytical control, and on trial-to-trail fluctuations of the arousal state and a reliable measure thereof. Below, I will first briefly describe these components. I will then summarize the results of an experiment that I conducted as part of my MSc studies, which forms the foundation of my PhD work. I will end the synopsis with a summary of each chapter, and a general discussion.

0.2 Decision-making under experimental and analytic control

A class of decisions that lends itself particularly well to scientific inquiry is that of the perceptual kind [2, 18–23]. Perceptual decision-making involves choosing one option from a set of alternatives based on the available sensory information. The goal is thus to interpret noisy sensory information and translate this into behavior. For example, in one task that I have used throughout this thesis (Figure 1A), subjects

decide if a degraded image contains only noise (black and white flickering dots, the same you see on a television when there is no signal), or if there is a weak target stimulus (oriented bars) embedded in the noise; they press one of two buttons to report their "yes" vs. "no" choice. Such a task might seem trivial. However, it conserves the key computation of a decision process: the accumulation of external evidence (in this case across unique frames of sensory information).

Perceptual decision-making is a powerful framework for four main reasons. First, perceptual decisions typically span only a few hundred milliseconds up to a few seconds. Subjects can thus make many decisions in one experimental session, which, in turn, means that one can compare robust measures of choice behavior (based on many trials) between relatively low and high arousal states.

Second, the quality of the evidence that the decision is based on is under complete experimental control. For example, in the contrast detection task (Figure 1A), the "signal-to-noise ratio" of the sensory evidence can be manipulated by changing the contrast of the signal (the oriented bars) while keeping the noise the same. That means that subject's performance can be fixed at a level that simulates decision-making under uncertainty (at about 75% correct). Fixing the available external evidence across repeated decisions also means that when the choice behavior is different between high and low arousal states, this difference should be attributed to internal factors.

Third, mathematical models of the assumed underlying cognitive processes responsible for perceptual decisions have been developed. That means that, by modeling the subject's overt behavior, a lot can be learned about how the brain's arousal state shapes the decision process. The work in this thesis relies heavily on signal detection theory [24] and the drift diffusion model [18, 25]. Signal Detection Theory decomposes subject's choice behavior into the orthogonal measures "sensitivity" (d') and "bias" (criterion) (Figure 1B). Sensitivity is the ability to discriminate signal from noise, and bias is the overall tendency to choose one of the choice alternatives over the other irrespective of the evidence. Signal Detection Theory can thus be used to test whether the brain's arousal state affects a person's sensitivity or bias. The drift diffusion model formally captures the idea that decision-making is a dynamic process that unfolds over time (Figure 1C). Specifically, it assumes that decisions are formed by continuously accumulating (sensory) information towards one of the two response boundaries (for "yes" and "no"); once such a bound is met, the decision process is terminated, and a motor response ensues. The model is fitted to subject's choices and reaction times, and can be used to test if the brain's arousal state affects any of the following components of the decision process: an initial bias towards one of the choice alternatives, the overall efficiency of evidence accumulation, a bias in evidence accumulation towards on the choice alternatives, how much evidence is required before one commits to a choice, or the speed of non-decisional processes like the pre-decisional evidence encoding and post-decisional translation of choice into a motor response. We can thus ask which of the *algorithmic* decision-computations are shaped by arousal, and how.

Fourth, the basic underlying physiological mechanisms that mediate perceptual decision-making in humans, monkeys and rodents are well understood [2, 18–23]. Specifically, neural activity in sensory cortex seems to lawfully depend on the signal-to-noise ratio of the external stimulus, and that trial-to-trial

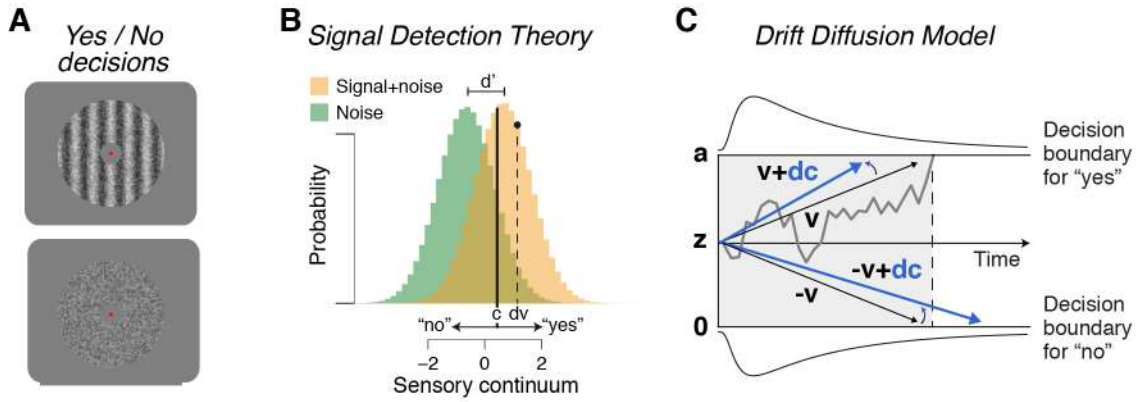


Figure 1: Perceptual decision-making and computational modeling of behavior. (A) Example stimulus used in perceptual decision-making. On each trial, either a signal+noise stimulus (top) is presented or only noise (bottom), and subjects indicate their "yes" or "no" choice with a button press. The signal is shown at high contrast for illustration purposes only. In an actual experiment, its contrast is typically lowered to a level that simulates decision-making in the face of uncertainty (yielding roughly 75% correct choices). **(B)** Signal detection theory. On each trial, the external stimulus (signal+noise or noise) is converted into an internal decision variable (dv). To make a binary choice, the observer compares the value of the decision variable to an internal choice criterion (c). Due to internal and external noise, a distribution of decision variables over trials can be modeled as a Gaussian, the variance of which reflects the amount of noise present in the system. Increasing the available sensory evidence or training will result in distributions that are more separable, and thus in better performance. By comparing the proportion of responses for the two external stimulus identities, signal detection theory can deduce the observers' perceptual sensitivity (d') and her choice criterion (c). **(C)** Schematic of drift diffusion model accounting for reaction time distributions for "yes" and "no" choices. In one example trial (grey), noisy sensory evidence is integrated over time into a decision variable, until reaching one of two choice bounds. The model consists of the following parameters: the starting point of evidence accumulation (an initial bias towards one of the choice alternatives, *z*), the evidence accumulation (the overall efficiency thereof, *v*, and a bias therein towards one of the choice alternatives, *dc*), boundary separation (controlling how much evidence is required before one commits to a choice, thereby implementing a speed-accuracy tradeoff, *a*) and the so-called non-decision time (the speed of pre-decisional evidence encoding and post-decisional translation of choice into motor response).

fluctuations in such stimulus-specific activity predicts choice. Together, this means that input signals to the decision-process ("evidence samples") are provided by sensory cortex. Likewise, association cortex seems to accumulate these input signals over time into a decision variable up to a bound, and again, these ramping signals lawfully depend on the signal-to-noise ratio of the external stimulus (steeper ramping for stronger stimuli) and predict choice. Note that the accumulation-to-bound framework of the drift diffusion model (described above) corresponds well to this biological decision architecture. Finally, motor regions seem to translate the decision into a behavioral act. Thus, perceptual decision-making is implemented by a distributed brain network (Figure 2). This is one of the reasons that each of these stages of decision-making can be measured non-invasively in the human brain with for example magnetoencephalography (MEG) [22, 26–28]. We can thus ask which of the *cortical* decision-computations are shaped by arousal, and how.

0.3 Pupillometry, a window into the brain's arousal state

Critically, the state of each cortical area in Figure 2 around the time of a decision is shaped by the interplay of cortical network interactions on the one hand, and neuromodulators released from ascending

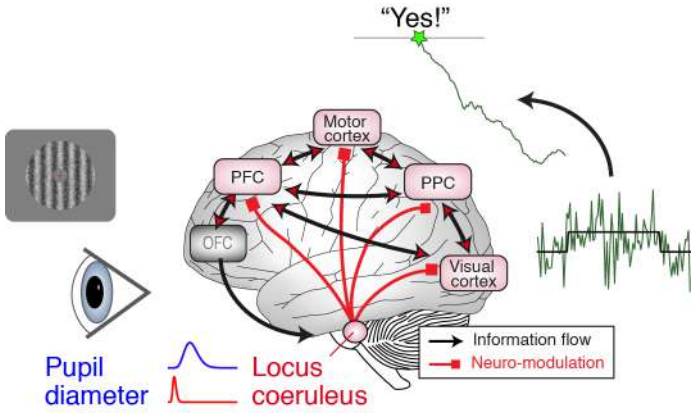


Figure 2: Neuromodulation of cortical decision dynamics. During visuo-motor decisions (e.g. based on contrast), "sensory evidence" is accumulated over time up to a bound, and translated into an action plan. This process unfolds in a recurrently connected network of fronto-parietal and motor cortices. Each node of the network is under the influence of ascending neuromodulatory brainstem systems, which exhibit phasic (fast) responses during decision formation. The modulatory neurotransmitters released by these systems regulate the brain's arousal state and have significant effects on the operating mode of local and large-scale cortical networks. Pupil size is a reliable read-out of phasic arousal. PPC, posterior parietal cortex; PFC, prefrontal cortex; OFC, orbital frontal cortex.

brainstem systems on the other hand (marked in red). Again, neuromodulatory brainstem nuclei have widespread projections to cortex and tune parameters governing the operating mode of their cortical target circuits, such as the "excitability" of synaptic interactions between different neurons [6–10]. An important recent insight is that neuromodulatory brainstem systems operate on multiple timescales, and with precise locking to cognitive events [6]. In particular, the noradrenergic locus coeruleus and cholinergic basal forebrain system can be rapidly recruited, in a time-locked fashion, during elementary decisions [6, 29–31]. Taken together, this means that the arousal state of the brain is briefly elevated every time we make a decision.

How we do track the brain's arousal state on each trial? As neuromodulatory nuclei are tiny and located deep within the brain, their activity is hard to measure with any neuroscientific method, and especially so with non-invasive techniques. Conveniently, it is now clear that pupil responses can be used as a peripheral readout of changes in cortical arousal state. The pupil size is most strongly driven by changes in the overall luminance. However, when light conditions are stable, the pupil still dilates and constricts, and these fluctuations track established measures of the brain's arousal state [11, 32]. Changes in pupil diameter are also associated with responses of the locus coeruleus in humans [33, 34], monkeys [35, 36], and mice [37–39].

Here it is important to highlight that 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 [33, 35, 39]. Thus, throughout this thesis, I remain agnostic about the exact neuroanatomical source(s) of the reported effects of transient (phasic) increases in the brain's arousal state during decision-making.

To test if and how arousal shapes decision-computations and choice behavior we require at least some

trial-to-trial variation in the magnitude of decision-related phasic arousal responses, because only in that case can choice behavior between relatively high and low arousal states be compared. As it turns out, indeed the strength of phasic arousal responses varies substantially from trial to trial. I discovered this in a pupillometry study [40] that I conducted during my MSc studies in the lab of Prof. Tobias Donner, my later PhD advisor. The results of this experiment allowed us to formulate specific hypotheses regarding the functional impact of phasic arousal, which I subsequently tested during my PhD project. I will therefore briefly summarize this work here.

0.4 Decision-related pupil dilation reflects upcoming choice and individual bias

Our first goal was to uncover at which point during a perceptual decision the pupil-linked phasic arousal signals occur. In a yes/no contrast detection task under constant luminance (as in Figure 1A), we indeed found that the pupil dilates time-locked to decisions. However, the relatively sluggish (~1 s; [41]) response dynamics of pupil dilation make it challenging to connect changes in pupil size to events occurring close together in time. For example, does the pupil dilate every time we start a decision-process, or is the pupil dilated for as long as we are forming a decision (during evidence accumulation), or does the pupil only dilate when we act on the decision that we reached? We used linear systems analysis to "de-mix" the pupil response timecourse and quantify the contribution of each of these three temporal components to task-evoked pupil responses. Task-evoked pupil responses were predominantly driven by a sustained input that took place from stimulus onset until the button press (Figure 3A,B) – in other words, throughout decision formation. This key result was later replicated by ourselves and others [33, 42, 43]. Given the prolonged nature of the decision (reaction times were typically ~1.5 s), this implies that cortical decision networks received a (neuromodulatory) arousal boost already *while* they were accumulation visual evidence towards "yes" or "no". This means that phasic arousal signals might indeed shape the decision-process and subsequent choice behavior.

Consistent with this idea, task-evoked pupil responses were larger for "yes" choices than for "no" choices, irrespective of whether these choices were correct (Figure 3C,D). Remarkably, the extent to which pupil responses were larger for "yes" than for "no" strongly depended on subjects' intrinsic bias (as measured in terms of criterion from signal detection theory, see also Figure 2B): it was amplified in subjects with strong biases (Figure 3E) In this dataset, subjects with strong biases tended to be conservative – that is, they tended to choose "no" irrespective of the objective evidence. These findings established a strong relationship between decision bias and phasic pupil-linked arousal.

Taken together, these results led us to hypothesize that phasic arousal shapes cortical decision-computations in a way that helps a decision maker to overcome his or her intrinsic bias. I tested this specific hypothesis over the subsequent years, the results of which are included in this thesis.

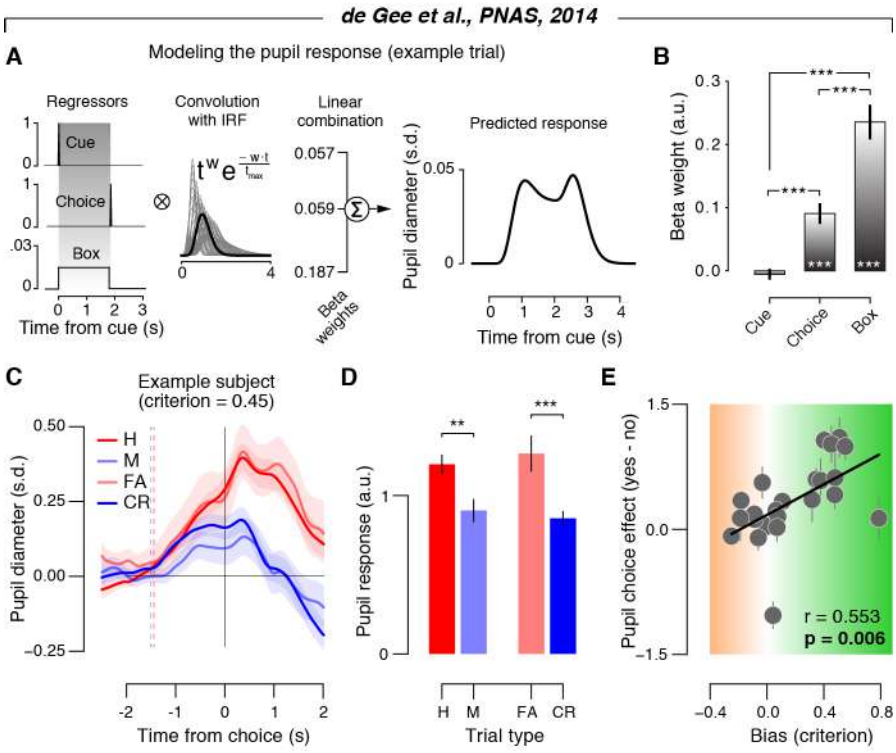


Figure 3: Decision related pupil dilation reflects upcoming choice and individual bias. (A) Illustration of general linear model to quantify temporal components driving task-evoked pupil responses. **(B)** Average of best-fitting beta weights for the three temporal components across all subjects (N = 23). **(C)** Decision-related pupil responses of one example subject, sorted according to signal detection theory. H (hit), "yes" choice on a signal+noise trial; M (miss), "no" choice on a signal+noise trial; FA (false alarm), "yes" choice on a noise trial; CR (correct reject), "no" choice on a noise trial. **(D)** Pupil response amplitudes ranked across the categories of signal detection theory as follows: hits = false alarms > misses = correct rejects. This gave rise to a strong choice-effect: more pupil dilation for "yes" than for "no" choices. **(E)** Pupil choice effect (average pupil response for "yes" minus "no" choices) per subject plotted against subjects' decision criteria, a measure of intrinsic bias. Adapted from [40].

0.5 Thesis overview and summary of results

The overarching goal of this thesis is to uncover how phasic arousal and the history of choices shape decision-computations (at the algorithmic and cortical levels) and subsequent choice behavior. I approached this question with a series of experiments in which I combined psychophysics, computational modeling, pupillometry, functional magnetic resonance imaging (fMRI) of the brainstem, and magnetoencephalography (MEG) of cortical decision computations.

Chapter 1 [33] establishes that task-evoked pupil responses in humans reliably track the task-evoked responses in a number of neuromodulatory brainstem nuclei known to regulate cortical arousal, including the noradrenergic locus coeruleus. Our study was the first (and remains the only till date) to quantify this relationship in humans. Further, the first chapter describes a behavioral correlate of phasic arousal: large pupil responses predict a reduction of suboptimal decision bias; computational modeling of behavior suggests that this is due to a reduction of bias in the evidence accumulation process. A strength of this

chapter is the concurrent brainstem fMRI and pupillometry during a simple contrast-dection task. This allowed us to characterize the relationship between the activity of neuromodulatory brainstem nuclei, pupil-linked phasic arousal and sophisticated (model-based) measures of subject's choice behavior.

Chapter 2 [44] demonstrates that the arousal-linked suppression of suboptimal biases in evidence accumulation occurs in different species (humans and mice) and in a range of decision-making tasks: when decisions are based on perceptual evidence, when evidence is accumulated from memory, or when decisions are based on numerical values. A strength of this chapter is the inclusion of six independent pupillometry datasets that involved evidence accumulation in various contexts and in different species. This allowed us to replicate the behavioral correlate of phasic arousal observed in chapter 1, and reveal that the arousal-linked suppression of suboptimal biases in evidence accumulation is a general principle.

Chapter 3 identifies the biological underpinnings of the pupil-linked bias suppression. In the context of a perceptual decision, the pupil-linked bias suppression seems to be mediated by the arousal-related shaping of choice-predictive alpha-band activity in regions along the dorsal stream of the visual cortical hierarchy, including primary visual cortex. A strength of this chapter is the concurrent MEG and pupillometry during a similar contrast-dection task as used in chapter 1. This allowed us to characterize *how* phasic arousal tunes cortical decision computations (that we could measure in real time) and subsequent behavior, and conclude that phasic arousal shapes evidence accumulation by tuning cortical excitability.

Chapter 4 [45] presents evidence for the idea that the trial-to-trial variation in the magnitude of pupil-linked phasic arousal responses is not random, but that instead phasic arousal is recruited in relation to key computational variables such as uncertainty and surprise. A strength of this chapter is the experimental design that allowed us to formally quantify uncertainty and surprise on each trial and observe their effects on arousal.

Chapter 5 [46] uncovers at the algorithmic level how previous choices influence current decisions. We found that individual history biases in overt behavior were consistently explained by a history-dependent change in the evidence accumulation, rather than in its starting point. Thus, evidence accumulation is not only driven by the objective evidence, intrinsic bias, and arousal, but also by previous choices. A strength of this chapter is the demonstration that this conclusion holds for various task protocols and sensory modalities and for various computational modeling approaches.

0.6 Discussion

Many decisions are made in the face of uncertainty and involve accumulating ambiguous evidence over time. This holds for elementary decisions based on perceptual evidence ranging to more complex decisions based on abstract evidence. Strikingly, when asked to make repeated decisions, we often choose

differently each time, even when we are given the same information to inform our choice [1–5]. This implies that the path a decision-making process takes in our brain is not fixed, but, to an extent, variable. Behavioral variability is commonly attributed to intrinsic noise in the nervous system [1, 3].

This thesis uncovers that a substantial part of the variability in behavior can be explained and predicted by tracking the momentary arousal state of the brain, and by tracking the history of choices a particular decision is embedded in. Key results include: pupil size at constant luminance is a reliable measure of the activity of a number of neuromodulatory brainstem nuclei known to regulate cortical arousal (chapter 1); the functional impact of pupil-linked phasic arousal is to suppress suboptimal biases in evidence accumulation and subsequent choice behavior (chapter 1); the arousal-linked suppression of suboptimal biases in evidence accumulation is a general principle: it occurs in a range of decision-making tasks (when decisions are based on perceptual evidence, are made from memory, or based on numerical values) and across species (humans and mice) (chapter 2). Pupil-linked bias suppression (in a perceptual decision-making context) is mediated by the arousal-related shaping of choice-predictive alpha-band activity (a signature of "excitability") in regions along the dorsal stream of the visual cortical hierarchy, including primary visual cortex (chapter 3); pupil-linked phasic arousal is recruited in relation to key computational variables such as uncertainty and surprise (chapter 4); previous choices shape the current evidence accumulation process, rather than its starting point (chapter 5). Thus, phasic arousal and previous choices calibrate a key computation during decision-making: evidence accumulation.

0.6.1 Post-decisional versus intra-decisional drive of phasic arousal

One "post-decisional" account holds that the phasic arousal signals (specifically, phasic responses of the locus coeruleus) are triggered by the bound-crossing in one of the cortical accumulator circuits; the resulting transient and cortex-wide neuromodulator release then facilitates the translation of the choice into a motor act [6]. This hypothesis is supported by single-unit physiology work in monkeys which showed that phasic locus coeruleus responses are more closely aligned to the behavioral response than to the onset of behaviorally relevant stimuli [47, 48] and that locus coeruleus responses encode sensory and motor events related to decisions to execute movements, but not to withhold movements [49]. An alternative "intra-decisional" idea [30], supported by indirect evidence [40, 42, 43], is that arousal systems are already recruited before the bound-crossing, throughout the evidence accumulation process. Within the drift diffusion model, the former notion predicts a reduction in non-decision time, while latter notion predicts an effect on evidence accumulation. In line with the intra-decision account, we found in all our datasets that phasic arousal affected evidence accumulation (suppressing biased therein), but not non-decision time (chapters 1–3). Lending further support to the intra-decision account, we observed pupil-linked alpha-band suppression (a MEG signature of neural excitability) already from stimulus onset onwards and more than a second before report – in other words, during evidence accumulation (chapter 3). These finding indicate that at least one of the brainstem nuclei regulating cortical arousal is activate during the deliberation process [20]. Because the tasks used in previous animal physiology studies of task-related lo-

cus coruleus responses involved much faster decision processes than the one studied here (reaction times of about 0.5 s vs. 2 s, respectively), it remains unknown whether the more sustained, task-evoked responses also occur in noradrenergic neurons (but see [36]). Future experiments characterizing phasic activity in the locus coeruleus or other brainstem nuclei involved in arousal during protracted evidence accumulation tasks could shed light on this issue.

0.6.2 How do diffuse arousal signals translate into specific effects on behavior?

Owing to their diffuse (widespread) projections, neuromodulatory brainstem nuclei are traditionally thought to exert uniform and synchronous modulatory effects on cells and circuits throughout the central nervous system [6]. How can such a non-specific signal cause the specific effect of reducing choice bias? There are a number of (non-mutually exclusive) scenarios consistent with that result.

A first possibility is that task-evoked neuromodulatory responses and cortical decision circuits interact in a recurrent fashion. This idea is line with a conceptual model that has been proposed for phasic locus coeruleus responses, which posits that cortical regions driving the locus coeruleus (e.g. anterior cingulate cortex) continuously compute the ratio of the posterior probability of the state of the world, divided by its (estimated) prior probability [30]. In other words, this framework predicts that the locus coeruleus is particularly activate when cortical decision-computations ramp up towards the choice that goes against one's bias (a conservative person about to report "yes"). The resulting locus coeruleus activity might reset its cortical target circuits [31] and override the default state [30], facilitating the transition of the cortical decision circuitry towards the non-default state.

A second possibility is that neuromodulators impinge on a cortical architecture that is spatially specific. Biophysically detailed models of the neural computations underlying decision-making [18] help to conceptualize this idea. In one specific scheme for yes/no choices, supported by neurophysiological data [26, 50, 51], two neural populations accumulate evidence for "yes" and "no" toward separate bounds, and compete via mutual inhibition [26, 52, 53]. The yes-population accumulates the sensory evidence for signal presence–i.e., neural activity in visual cortex, which is either stimulus induced (on signal+noise trials) or spontaneous (on noise trials) [54]. The no-population accumulates a "default input" [52]. If, for example, visual cortex has more receptors for the neuromodulators to bind to than the region providing the "default input", then boosting global arousal might asymmetrically enhance neural activity in visual cortex (on both signal+noise and noise trials). This, in turn, will translate in faster accumulation toward "yes" and the reduction of a conservative bias. Consistent with this idea, G protein-coupled receptors indeed have a distinct and heterogeneous distribution across cortical areas [55], which partly explains the spatially structured effects of noradrenaline and dopamine on the forebrain [56].

A third (related) possibility is that neuromodulatory responses alter the balance between "bottom-up" and "top-down" signaling across the cortical hierarchy [57–61]. Sensory cortical regions encode likelihood signals and send these (bottom-up) to association cortex; participant's prior beliefs (about for example target probability) are sent back (top-down) to the lower levels of the hierarchy [57, 62]. Thus, through a

relative suppression of top-down signal flow, phasic arousal might reduce the weight of the prior (reflecting subjects' intrinsic bias) relative to the likelihood. Specifically, in the yes/no task used throughout this thesis, the prior may have been a conservative bias for choosing "no". Reducing its weight would reduce this bias.

A fourth (related) possibility is the ascending projections from for example the cholinergic basal forebrain and locus coeruleus are more spatially specific than commonly believed. Consistent with that idea, locus coeruleus cells innervating discrete cortical regions are biochemically and electrophysiologically distinct from one another [63–66].

0.6.3 Conclusion

Phasic arousal and previous choices calibrate a key computation during decision-making: evidence accumulation. When decisions are made in the face of uncertainty tracking phasic arousal signals and the history of choices may be just as important for predicting choice behavior as tracking the objective evidence gathered from the outside world. This thesis therefore establishes that phasic arousal signals and the history of choices account for a substantial fraction of the apparent variability in overt behavior.

Status: published
de Gee JW, Colizoli O, Kloosterman N, Knapen T, Nieuwenhuis S & Donner TH (2017). *eLife,* 6, e23232.

1

Dynamic modulation of decision biases by brainstem arousal systems

Conceptualization by JWG, THD; Investigation by JWG, OC, NAK; Formal analysis by JWG; Resources by TK, SN, THD; Writing–original draft by JWG, THD; Writing–review and editing by JWG, OC, NAK, TK, SN, THD.

Decision-makers often arrive at different choices when faced with repeated presentations of the same evidence. Variability of behavior is commonly attributed to noise in the brain's decision-making machinery. We hypothesized that phasic responses of brainstem arousal systems are a significant source of this variability. We tracked pupil responses (a proxy of phasic arousal) during sensory-motor decisions in humans, across different sensory modalities and task protocols. Large pupil responses generally predicted a reduction in decision bias. Using fMRI, we showed that the pupil-linked bias reduction was (i) accompanied by a modulation of choice- encoding pattern signals in parietal and prefrontal cortex and (ii) predicted by phasic, pupil-linked responses of a number of neuromodulatory brainstem centers involved in the control of cortical arousal state, including the noradrenergic locus coeruleus. We conclude that phasic arousal suppresses decision bias on a trial-by-trial basis, thus accounting for a significant component of the variability of choice behavior.

Decision-makers often arrive at different choices in the face of repeated presentations of the same evidence [1, 2, 4, 5]. This intrinsic behavioral variability is typically attributed to spontaneous fluctuations of neural activity in the brain regions computing decisions [1, 3] (but see [67, 68]). Indeed, fluctuations of neural activity are ubiquitous in the cerebral cortex [1, 69, 70].

One candidate source of these fluctuations in cortical activity is systematic variation in central arousal state. Central arousal state is controlled by the neuromodulatory systems of the brainstem, which have widespread projections to cortex and tune neuronal parameters governing the operating mode of their cortical target circuits [6, 8, 9]. Importantly, these neuromodulatory systems operate at different timescales [6, 29]. Some, in particular the noradrenergic locus coeruleus (LC), are rapidly recruited, in a time-locked fashion, during elementary decisions [6, 29–31]. Pupil diameter, a reliable peripheral marker of central (cortical) arousal state [11], also increases during decisions [40, 71–74]. These observations point to an important role of phasic (i.e., fast) pupil-linked arousal signals in decision-making [6, 30]. Yet, the precise nature of this role has remained unknown.

Here, we investigated how phasic, task-related arousal interacts with decision computations in the human brain. We combined pupillometry, fMRI, and computational modeling to probe into the interplay between task-related arousal and decision computations underlying elementary sensory-motor choice tasks. Sensory-motor decisions entail the gradual accumulation of noisy "sensory evidence" about the state of the world towards categorical decision states governing behavioral choice [2, 18, 19, 25]. A large-scale network of regions in frontal and parietal cortex seems to accumulate stimulus responses provided by sensory cortices towards choices of motor movements [2, 22] (but see [19, 75]). We here aimed to elucidate the interaction between pupil-linked arousal responses, evidence accumulation, and decision processing across several (cortical and subcortical) brain regions.

Large task-evoked pupil responses were consistently accompanied by a reduction in perceptual decision bias in different sensory modalities (visual and auditory) and task protocols (detection and discrimination). Decision bias reflects the degree to which an observer's choice deviates from the objective sensory evidence. Using fMRI for one of these tasks revealed that the bias reduction was accompanied by a modulation of choice-encoding pattern signals in prefrontal and parietal cortex. Further, the bias reduction was predicted by task-evoked, pupil-linked responses in a network of neuromodulatory brainstem nuclei controlling cortical arousal state. We conclude that phasic neuromodulatory signals reduce biases in the brain's decision-making machinery. As a consequence, phasic arousal accounts for a significant component of the variability of choice behavior, over and above the objective evidence gathered from the outside world.

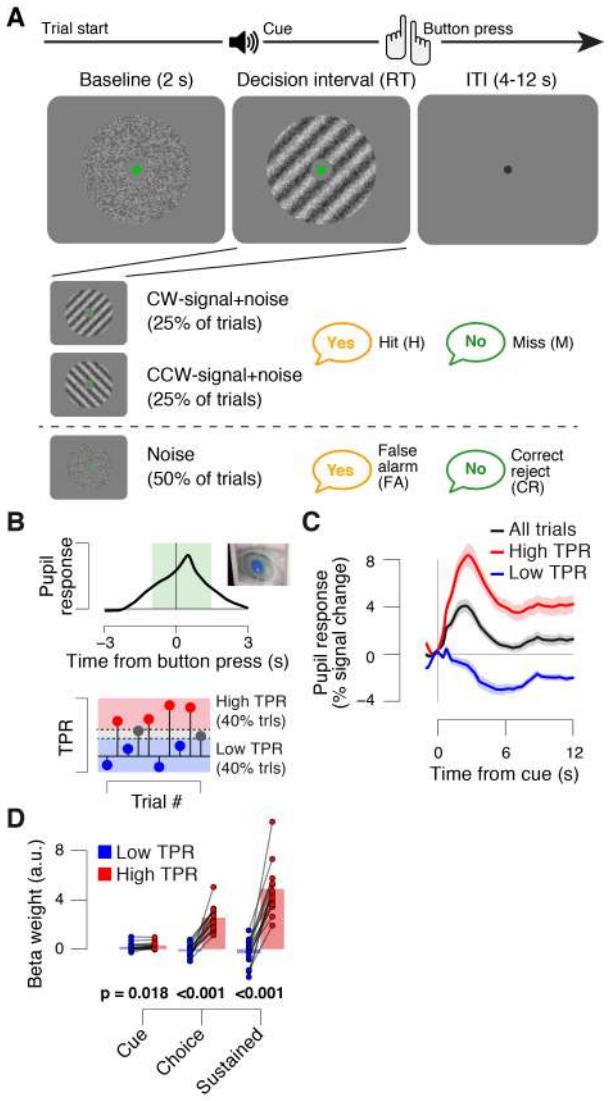


Figure 1: Behavioral task and task-evoked pupil responses. (A) Yes-no contrast detection task. Top: schematic sequence of events during a signal+noise trial. Subjects reported the presence or absence of a faint grating signal superimposed onto dynamic noise. Bottom left: the signal, if present, was oriented clockwise or counter clockwise on different blocks (known to the subject beforehand). Signal contrast is high for illustration only. Bottom right: trial types. **(B)** Quantifying task-evoked pupillary response (TPR) amplitude. Top: mean TPR time course of an example subject. Green box, interval for averaging TPR values on single trials. Bottom: trials were pooled into three bins of TPR amplitudes (lowest/highest 40% and intermediate 20%). **(C)** TPR time course for the three bins. **(D)** Mean beta weights of transient (cue, choice) and sustained input components under low vs. high TPR, estimated with a general linear model (see Materials and methods; Figure 1–figure supplement 1A,B), separately for low and high TPR trials. Panels C, D: group average (N = 14); shading, s.e.m.; data points, individual subjects; stats, permutation test.

1.2 Results

We systematically quantified the interaction between pupil-linked arousal responses and decision computations at the algorithmic and neural levels of analysis. We here operationalize "phasic arousal" as task-evoked pupil responses (TPR). This operational definition is based on recent animal work, which established remarkably strong correlations between non-luminance mediated variations in pupil diameter and

global cortical arousal state [11].

The Results section is organized as follows. First, we quantify TPRs during the main behavioral task studied in this paper. The key observation here was the substantial trial-to-trial variability of the TPR amplitude. All subsequent analyses exploited this variability to pinpoint the functional correlates of phasic arousal. We then present results from modeling TPR-dependent changes in choice behavior, identifying precise algorithmic correlates of phasic arousal. These results yielded detailed predictions for the underlying modulations of cortical signals. Third, we present tests of these predictions, focusing on functionally delineated cortical regions of interest. We conclude by establishing that the trial-to-trial fluctuations in TPR amplitude, and the associated bias reduction, were closely linked to task-evoked responses of neuromodulatory brainstem centers involved in regulating cortical arousal state.

1.2.1 Tracking trial-to-trial fluctuations in phasic arousal

The main task used in this study was detection ("yes-no", simple forced choice protocol) of a low-contrast grating (Figure 1A). The grating contrast was titrated to the 75% correct level, and subjects did not receive trial-by-trial feedback. As observed previously [40], TPR amplitudes during this task fluctuated widely from trial to trial (Figure 1B,C; see Materials and Methods for quantification of TPR). To illustrate, pooling trials into two bins containing the lowest and highest 40% of TPR amplitudes (Figure 1B) yielded, on average, the commonly observed task-evoked pupil dilations for the high TPR bin, but pupil constrictions for the low TPR bin (Figure 1C). We used a previously established model to estimate the time course of the neural input driving the measured TPRs (GLM; see Materials and methods; Figure 1–figure supplement 1A–C). This revealed that the difference between the low and high TPR bins was primarily due to the difference in a sustained component that spanned the entire interval from cue to behavioral choice (Figure 1D). The difference of the sustained component between low and high TPR was significantly larger than the corresponding difference for two components at cue or choice, respectively (2-way repeated measures ANOVA with factors temporal component and TPR bin; interaction: $F_{2,26} = 79.00$, $p<0.001$).

In sum, TPR amplitude exhibited substantial trial-to-trial fluctuations, which were predominantly driven by changing levels in sustained input during decision formation. Given the prolonged nature of the decision (median of subject-median reaction time, RT: 2.11 s), the sustained, intra-decisional arousal boost might have interacted with the decision computation. To test for such an interaction between arousal boost and decision computation, we next modeled subjects' choice behavior as a function of TPR amplitude.

1.2.2 Phasic arousal is inversely related to decision bias

We found a robust and consistent relationship between TPR and decision bias. This effect was present in two independent data sets using an analogous contrast detection task: the newly collected fMRI data set, and a re-analysis of an existing data set [40]) (Figure 2A,D, middle and right panels). Decision bias was

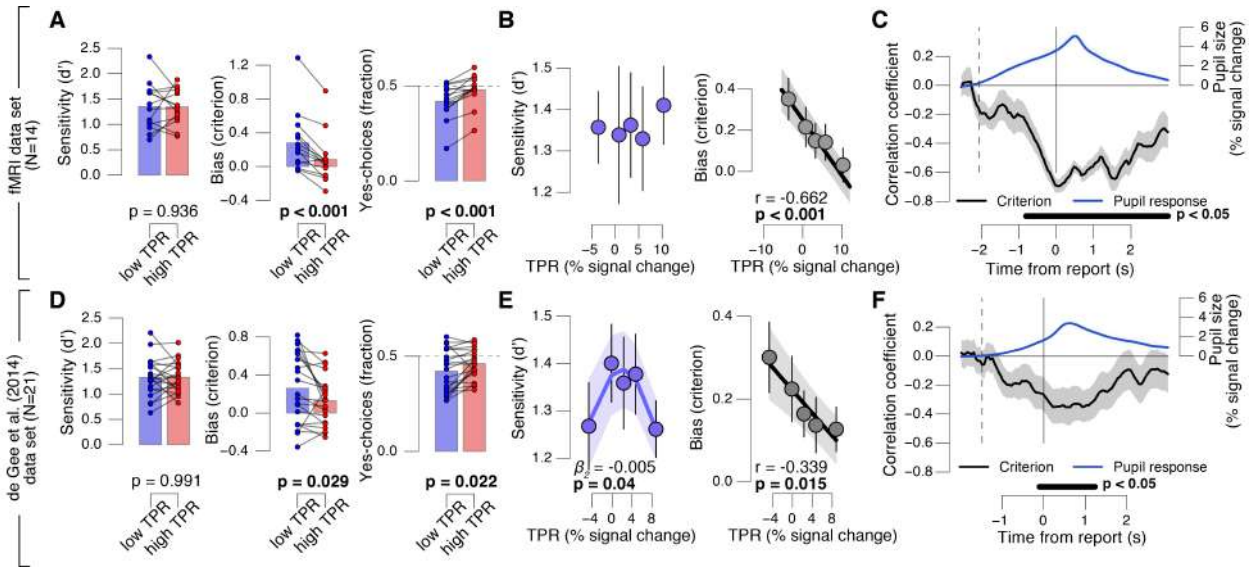


Figure 2: Phasic arousal predicts reduction of choice bias. (A) Perceptual sensitivity SDT d' (left), decision bias, measured as SDT criterion (middle) or fraction of "yes"-choices (right), for low and high TPR. For the fraction of "yes"-choices analysis, we ensured that each TPR bin consisted of an equal number of signal+noise and noise trials (see Materials and methods). Data points, individual subjects. **(B)** Relationship between TPR and d' or criterion (5 bins). Linear fits are plotted wherever the first-order fit was superior to the constant fit (see Materials and methods). Quadratic fits are plotted wherever the second-order fit was superior to first-order fit. **(C)** Sliding window linear correlation between TPR and SDT criterion (5 bins), aligned to button press. Dashed line, median decision onset (cue). The group average pupil response time course is plotted for reference in blue. (D–F) As panels A-C, for an independent data set [40]. All panels: group average (N = 14 and N = 21); shading or error bars, s.e.m.; stats, permutation test.

quantified in two ways (for details, see Materials and methods). First, we computed signal detection-theoretic (SDT) criterion (Figure 2A,D, middle panels). Second, we computed the fraction of "yes"-choices (right panels), after balancing the number of signal+noise and noise trials within each TPR bin. We did not find a consistent relationship between phasic arousal, as measured by TPR, and perceptual sensitivity, quantified by SDT d' (Figure 2A,D, left panels).

The negative association between TPR and decision bias (SDT criterion) was approximately linear across a range of five TPR-defined bins (Figure 2B,E, right panels). In all cases, here and below, we tested whether fits of second-order polynomials, reflecting non-monotonic relationships between TPR and behavior, were superior to the linear fits (via sequential polynomial regression analysis; Materials and methods). We found a non-monotonic relationship between TPR and sensitivity in the behavioral data set from [40], but not in the fMRI dataset (Figure 2B,E, left panels).

This non-monotonic (inverted U-shape) relationship between pupil diameter and sensitivity is consistent with previous animal work on correlations between baseline arousal and behavior [6, 76]. However, it was less consistent across the data sets analyzed in this paper than the negative linear effect of TPR on decision bias. The consistent effect of TPR on decision bias has not been reported before in previous studies of slow fluctuations of baseline pupil diameter. In what follows, we focus on the negative effect of TPR on decision bias.

Most subjects were overall (i.e., without splitting trials by TPR) intrinsically biased to respond "no":

10 out of 14 subjects exhibited a significantly conservative criterion (within-subject permutation tests; $p<0.05$) in the fMRI data set, and 14 out of 21 subjects in the data set from [40]. Because signal+noise and noise trials were equally frequent in both experiments, this bias was always maladaptive. Critically, this maladaptive bias was particularly pronounced under low TPR; but under high TPR the bias was nearly neutralized, especially in the fMRI data set (criterion around zero, and fraction of "yes"-choices around 0.5 for highest TPR bins, Figure 2A,B).

1.2.3 A robust effect of phasic arousal on the decision computation

A number of control analyses and experiments supported the idea that the negative correlation between TPR amplitude and decision bias reflected a specific effect of phasic arousal on the decision computation that generalized across perceptual choice tasks. First, the effect emerged during, not after, decision formation: a sliding-window correlation between TPR and criterion became negative from decision onset onwards, and reached statistical significance before button press (Figure 2C,F). In the fMRI data set, this correlation was highly significant more than 800 ms before button press (Figure 2C). Given the sluggish nature of the pupil response (see above), the underlying central arousal transients must have occurred even earlier than that, leaving substantial time for shaping the decision outcome.

Second, there was no robust association between baseline pupil diameter and decision bias (Figure 2–figure supplement 1A–D). This ruled out possible concerns that the effect might be due to corresponding (opposite) associations between baseline pupil diameter and behavior, "inherited" by TPR through its negative correlation with baseline pupil diameter [40].

Third, the effect of TPR on decision bias was robust with respect to the details of the analysis approach. For Figure 2, as for all other analyses reported in the main text, we removed (via linear regression) components explained by RT. The rationale was to specifically isolate variations in the amplitudes of the neural responses driving TPR, irrespective of RT, variations of which might also cause variations of TPR amplitude without changes in the underlying neural response amplitudes (for details see Materials and methods). We observed the same linear effect of TPR on bias without removing trial-to-trial variations in TPR that were due to RT (Figure 2–figure supplement 1E–J).

1.2.4 Pupil-linked bias reduction is a general phenomenon

Fourth, the effect of TPR on decision bias shown in Figure 2 generalized to other perceptual choice tasks, which differed on several dimensions from the main contrast detection task used in this paper (Figure 3). In one follow-up experiment, we measured pupil-linked behavior during an auditory yes-no (tone-in-noise) detection task near psychophysical threshold using the same stimuli as in [76] (see Materials and methods). The only visual stimulus was a stable fixation dot. The decision interval contained only auditory noise (the same as in [76]) on half the trials, and a pure sine wave superimposed onto the noise on the other half of the trials. Again, TPR predicted a significant (linear) reduction in conservative decision

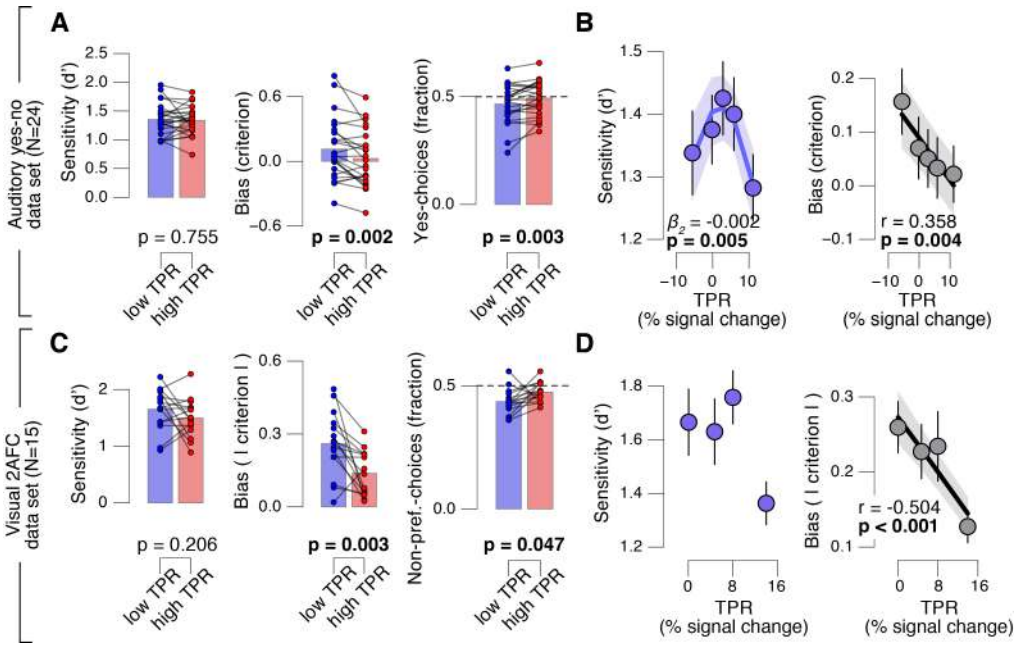


Figure 3: Arousal-linked bias reduction generalizes to other choice tasks. (A) Perceptual sensitivity (d'; left) and decision bias, measured as criterion (middle) or fraction of "yes"-choices (computed as for Figure 2A, right), for low and high TPR. Data points, individual subjects. **(B)** Relationship between TPR and d' or criterion (5 bins). Linear fits were plotted wherever the first-order fit was superior to the constant fit (see Materials and methods). Quadratic fits were plotted wherever the second-order fit was superior to first-order fit. **(C)** Perceptual sensitivity (d', left) and decision bias, measured as absolute criterion (middle) or fraction of non-preferred choices (right), for low and high TPR. For the fraction of non-preferred choices analysis, we ensured that each TPR bin consisted of an equal number of motion up and down trials (see Materials and methods). **(D)** Relationship between TPR and d' or absolute criterion (4 bins instead of 5, because of fewer trials per subject, see Materials and methods). All panels: group average (N = 24 and N = 15); shading or error bars, s.e.m.; stats, permutation test.

bias, and an increased tendency to respond "yes" (Figure 3A,B). TPR also exhibited a non-monotonic relationship with sensitivity, as observed in rodents for baseline pupil diameter in [76].

Another follow-up experiment assessed whether the pupil-linked bias reduction observed above may have been due to the asymmetric nature of the detection tasks (i.e., discriminating the presence from the absence of a signal) or due to the absence of single-trial feedback. Symmetric two-alternative forced choice tasks are commonly associated with weaker biases than yes-no detection tasks [24]. We used a symmetric visual random dot motion (up vs. down) discrimination task near psychophysical threshold with feedback after each trial (see Materials and methods). Although many subjects exhibited clear biases for reporting one or the other direction, these were more evenly distributed around zero than in the above yes-no tasks, in which the sign of the bias was largely consistent across individuals. Therefore, we here analyzed subjects' absolute criterion values (i.e., overall bias regardless of sign) and fraction of non-preferred choices (i.e., the choice opposite to their general bias, irrespective of TPR). Again, TPR predicted a reduction in absolute decision bias, and an increase in the fraction of non-preferred choices (Figure 3C,D), analogous to the effects observed for the detection tasks above.

In sum, a number of analyses and experiments showed that pupil-linked, phasic arousal was consistently associated with a monotonic reduction in perceptual decision biases in different sensory modalities

and task protocols.

1.2.5 Phasic arousal predicts a reduction of evidence accumulation bias

To further pinpoint the nature of the TPR-induced bias suppression, we fitted the drift diffusion model, an established dynamic model of two-choice decision processes (Figure 4A; [25]) to subjects' RT distributions from the main task (contrast detection). The drift diffusion model posits the perfect accumulation of noisy sensory evidence towards one of two decision bounds, here for "yes" and "no" (Figure 4A).

We fitted the model separately for low and high TPR trials (see Figure 4B for an individual example). Within the model, the TPR-induced reduction of conservative bias, evident in Figures 2 and 3, may have been brought about by two distinct mechanistic scenarios: (i) the evidence accumulation process started from a level closer to the "yes"-bound (i.e., a change in the "starting point" parameter); or (ii) the accumulation process was driven more towards the "yes"-bound (i.e., a change in the "drift criterion" parameter). The drift criterion is equivalent to an evidence-independent constant added to the drift. A non-zero drift criterion results in a bias of the decision variable that grows linearly with time. Although clearly distinct in nature, both mechanisms (starting point and drift criterion) would have resulted in an increase in the fraction of "yes"-choices, and thus a reduction of decision bias. Critically, both mechanisms were distinguishable through their distinct effects on the shape of the RT distribution (Figure 4–figure supplement 1). To dissociate between these alternative mechanisms we fitted the model, while allowing several model parameters (boundary separation, non-decision time, mean drift rate, starting point, and drift criterion) to vary with TPR.

The model fits (see Materials and methods and [77]) supported the second mechanism: a change in drift criterion. An individual example is shown in Figure 4B, and group data are shown in Figure 4C. Drift criterion was generally negative, indicating an overall conservative accumulation bias towards the bound for "no"-choices. But drift criterion was pushed closer towards zero under high TPR, indicating an unbiased drift, as optimal for the current task (Figure 4B,C). The other main parameters (including starting point and mean drift rate) were not significantly affected by TPR. The TPR-linked effect on drift criterion was also evident in the individual point estimates from the fMRI sample only (Figure 4D).

Again, we we found no evidence for an effect on any parameter of the drift diffusion model when comparing trials with low and high baseline pupil diameters (Figure 4–figure supplement 2A), and we obtained qualitatively identical results without removing trial-to-trial variations of RT from the TPR amplitudes (Figure 4–figure supplement 2B–D; Materials and methods).

As a control of the significance of the TPR-dependent effect on drift criterion, we re-fitted the model, but now fixing drift criterion with TPR, while still allowing all other of the above parameters to vary with TPR. In this variant of the model, we again found no TPR-dependent change in any of the other parameters (boundary separation: $p=0.428$; non-decision time: $p=0.370$; starting point: $p=0.117$; mean drift rate: $p=0.361$). Critically, model comparison favored the complete version of the model with TPR-dependent variation in drift criterion (deviance information criterion, 50437 vs. 50528, respectively; see Materials and

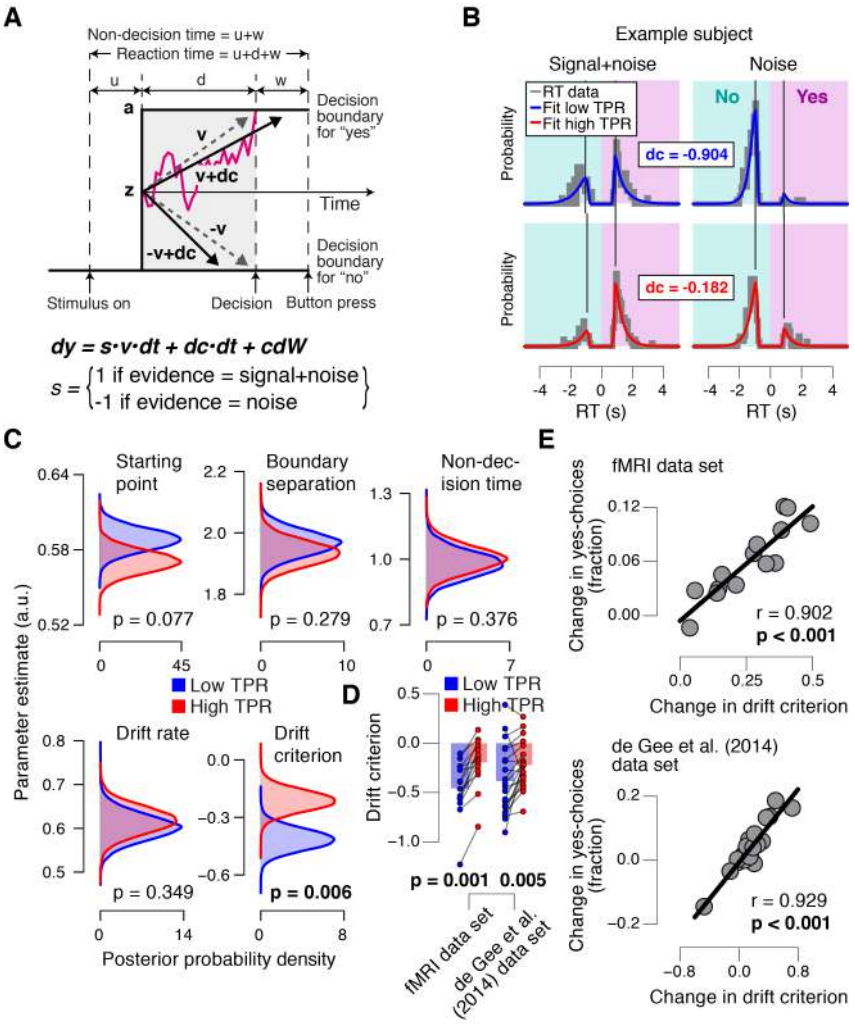


Figure 4: Phasic arousal predicts reduction of accumulation bias. (A) Schematic and simplified equation of drift diffusion model accounting for RT distributions for "yes"- and "no"-choices ("stimulus coding"; see Materials and methods). Notation: *dy*, change in decision variable y per unit time *dt*; $v \times dt$, mean drift (multiplied with 1 for signal+noise trials, and −1 for noise trials); $dc \times dt$, drift criterion (an evidence-independent constant added to the drift); and *cdW*, Gaussian white noise (mean = 0, variance = c^2 dt). **(B)** RT distributions of one example subject for "yes"- and "no"-choices, separately for signal+noise and noise trials and separately for low and high TPR. RTs for "no"- choices were sign-flipped for illustration purposes. Straight lines, mode (i.e., maximum) of the fitted RT distributions. Please note that TPR predicts an increased fraction of "yes"-choices with only a minor change of the mode of the RT distribution, consistent with a drift criterion effect rather than a starting point effect (Figure 4–figure supplement 1). **(C)** Group-level posterior probability densities for means of parameters. To maximize the robustness of parameter estimates [77], two data sets were fit jointly (the current fMRI and our previous study [40]; N = 35). Starting point (z) is expressed as a proportion of the boundary separation (a). **(D)** Drift criterion point estimates for low and high TPR trials, separately for both data sets (N = 14 and N = 21, respectively). Data points, individual subjects; stats, permutation test. **(E)** Change in fraction of "yes"- choices for low vs. high TPR trials, plotted against change in drift criterion. Data points, individual subjects.

methods). This implies that the TPR-dependent variability in accumulation bias was essential to account for the TPR-dependent effects on behavior.

The individual changes in drift criterion between low vs. high TPR trials established by means of diffusion modeling accounted for a substantial fraction of the individual differences in TPR-predicted changes in the fraction of "yes"-choices (Figure 4E) obtained in the model-free analyses (Figure 2A, D,

right panels). TPR-related changes in starting point had a weaker, and statistically not significant, effect on the fraction of "yes"-choices (fMRI data set: r=0.345, p=0.227; de Gee et al. (2014) data set [40]: r=0.419, p=0.059).

In sum, in the decision task studied here, pupil-linked, phasic arousal predicted a reduction of conservative bias, specifically in the evidence accumulation, and was neither reflected in the baseline level of the decision variable at the start of the accumulation nor its mean drift. In other words, TPR accounted for a portion of the trial-to-trial variability in the drift unrelated to the objective sensory evidence. This correlate of phasic arousal at the algorithmic level was in line with the notion that phasic arousal shapes decision outcome by interacting with the evidence accumulation computation that lies at the heart of the decision process.

Taken together, the behavioral modeling results reported in Figures 2–4 put strong constraints on the expected changes in cortical decision processing due to phasic arousal. Specifically, changes in the encoding of the incoming evidence by sensory cortical areas, as observed in previous work on fluctuations in baseline arousal levels [76, 78, 79], would be associated with changes in perceptual sensitivity. However, we found that TPR was not associated with any robust change in sensitivity (measured as d' or as mean drift rate) in the fMRI dataset, thus, predicting no TPR-linked modulation of sensory responses in visual cortex. Instead, the observed effect of TPR on choice bias (criterion, drift criterion) predicted a directed shift (towards "yes") in neural signals encoding subjects' choices, in downstream cortical regions. We next tested these predictions by assessing the relationship between TPR and (i) stimulus-specific responses in early visual cortex, and (ii) choice-specific responses in downstream cortical regions.

1.2.6 Phasic arousal does not boost sensory responses in visual cortex

The fMRI response in early visual cortex (areas V1, V2, and V3) during near-threshold visual tasks is made up of distinct components, including a (weak and focal) stimulus-specific component and a (large and global) task-related, but stimulus-independent, component [80–82]. We used an approach based on multi-voxel pattern analysis analogous to previous work [83, 84] to isolate the stimulus-specific response component. Because the majority of visual cortical neurons encoding stimulus contrast are also tuned to stimulus orientation, orientation-tuning could serve as a "filter" to separate the cortical stimulus response from stimulus-unrelated signals. Specifically, the low contrast signal in our task should have evoked a small response in each visual cortical neuron selective for the orientation of the target signal (45°or 135°, on different experimental runs, Figure 1A) across a substantial part of the retinotopic map. Thus, the presence or absence of the target signal should be reliably encoded in the orientation-specific component of the cortical population response, within the retinotopic sub-region corresponding to the signal. We first individually delineated these retinotopic sub-regions within each of V1-V3 (see Figure 5A for an example subject) and then quantified the orientation-specific response component therein as the spatial correlation of multi-voxel response patterns with an orientation-specific "template" (Materials and methods).

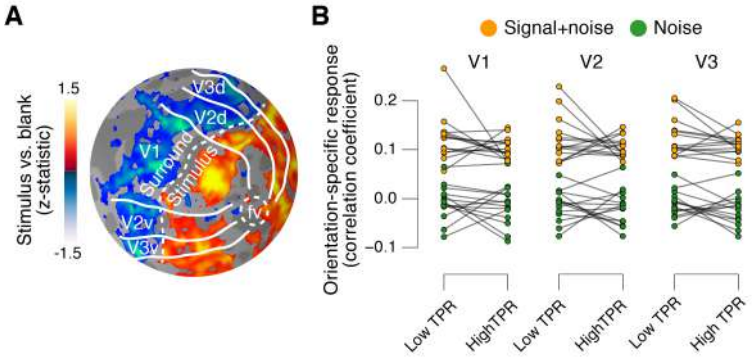


Figure 5: Phasic arousal does not boost sensory responses in visual cortex. (A) Map of fMRI responses during stimulus localizer runs (see Materials and methods); example subject. V1-V3 borders were defined based on a separate retinotopic mapping session. "Stimulus sub-regions", regions with positive stimulus-evoked response; "surround sub-regions", regions with negative stimulus-evoked response. **(B)** Orientation-specific fMRI responses in "center" sub-regions of V1-V3, separately for signal+noise and noise trials, and separately for low and high TPR trials. Statistical tests are reported in main text. Data points, individual subjects (N = 14); stats in main text.

As expected, this orientation-specific response component differed robustly between signal+noise and noise trials (Figure 5B). A 2-way repeated measures ANOVA with factors stimulus and TPR bin yielded a highly significant main effect of stimulus for V1, V2, and V3 (V1: $F_{1,13} = 303.5$, V2: $F_{1,13} = 646.3$, V3: $F_{1,13} = 316.6$; all $p<0.001$).

The orientation-specific response component also reliably discriminated between signal+noise and noise trials on a single-trial basis (Figure 5–figure supplement 1). Consequently, we henceforth refer to this component as the "stimulus-specific response". However, the stimulus-specific response was not boosted under high TPR (Figure 5B, no significant main effect of TPR, nor stimulus × TPR interaction in any of V1-V3).

1.2.7 No evidence for arousal-dependent boost of sensory responses in any cortical area

The above analysis focused on the stimulus-specific response in early visual cortex. To avoid missing TPR-dependent modulations of sensory responses in higher cortical regions, we also mapped out modulations of fMRI responses by TPR across cortex (see Materials and methods). Various regions including visual, parietal, prefrontal, and motor cortices exhibited robust task-evoked overall fMRI responses (i.e., difference between the decision interval and baseline; Figure 6A), as well as robust modulations by TPR (Figure 6B), whereby TPR-induced boosts only partly overlapped with the task-positive responses.

However, in no single region did the overall fMRI responses differ between signal+noise and noise trials (Figure 6C). This indicates that our multi-voxel pattern approach described above was, in fact, essential for detecting the weak cortical response to the near-threshold target signals. Critically, in no region did we find a significant interaction between the factors stimulus (signal+noise vs. noise) and TPR (low vs. high TPR; Figure 6D).

Taken together, both complementary analyses showed that phasic, task-evoked arousal signals did not

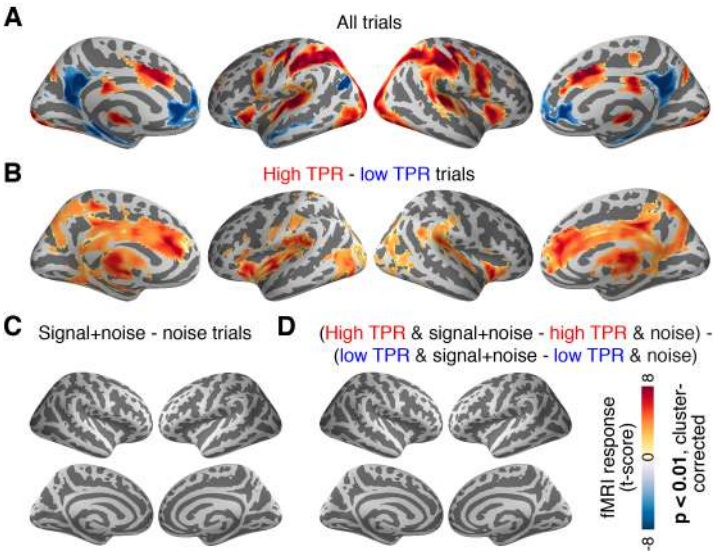


Figure 6: Cortex-wide fMRI correlates of phasic arousal and stimulus. (A) Functional map of task-evoked fMRI responses computed as the mean across all trials. **(B)** As panel A, but for the contrast high vs. low TPR trials. **(C)** As panel A, but for the contrast signal+noise vs. noise. **(D)** As panel A, but for the interaction between TPR (2 levels) and stimulus (2 levels). All panels: functional maps are expressed as t-scores computed at the group level (N = 14) and presented with cluster-corrected statistical threshold (see Materials and methods).

modulate cortical responses encoding the presence of the low-contrast signal. This is in line with the lack of TPR-linked change in perceptual sensitivity in the fMRI dataset (Figure 2A, Figure 4D).

1.2.8 Phasic arousal modulates choice-specific signals in frontal and parietal cortex

We then sought to test for directed shifts in neural signals encoding subjects' choices under high TPR, which would be in line with the changes in decision biases identified by behavioral modeling. Here, we use the term "choice-specific" to refer to fMRI-signals that reliably discriminated between subjects' choice ("yes" vs. "no"). Two complementary approaches delineated several cortical regions that exhibited such choice-specific signals (Figure 7). The first approach (Figure 7A) was based on the lateralization of fMRI responses with respect to the motor effector used to report the choice (i. e., response hand; see [26, 85] and Materials and methods). In addition to the hand area of primary motor cortex (henceforth referred to as M1), this approach yielded reliable effector-specific lateralization also in two regions of posterior parietal association cortex: the junction of the intraparietal and postcentral sulcus (IPS/PostCeS) and the anterior intraparietal sulcus (aIPS1; Figure 7A and Figure 7–figure supplement 1A,B). The second approach (Figure 7B) was based on multi-voxel pattern classification of choice, using a "searchlight" procedure that scanned the entire cortex for choice information (see [86, 87] and Materials and methods). The underlying rationale was to identify cortical regions encoding choice in other formats (e.g., in terms of more fine-grained patterns) than the hemispheric lateralization of response amplitudes. The second approach revealed robust (and reproducible) choice-specific response patterns in a number of additional regions in bilateral posterior parietal cortex and (right) prefrontal cortex: superior and inferior parietal lobule (SPL

and IPL, respectively), a second region within aIPS (aIPS2), posterior insula (pIns), the junction of pre-central sulcus and right inferior frontal gyrus (PreCeS/IFG) and right medial frontal gyrus (MFG; Figure 7B and Figure 7–figure supplement 1C, D). In both approaches, choice specific regions were delineated after factoring out the physical stimulus (see Materials and methods).

In all the above choice-encoding regions, responses (estimated in a cross-validated fashion, see Materials and methods) reliably differentiated between "yes"- and "no"-choices — both on average (Figure 7–figure supplement 1E,F) and at the single-trial level (Figure 7C, see also Figure 7–figure supplement 1G). As expected, the single-trial reliability of the choice-specific responses differed between cortical regions (1-way repeated measures ANOVA with factor region of interest (9 levels): $F_{8,104} = 30.20$, $p<0.001$), with the strongest reliability for M1 (dashed horizontal line in Figure 7C), the region closest to the subjects' motor output.

For analysis of the association with TPR, we pooled the choice-specific signals of these different regions into three groups (Figure 7–figure supplement 1A): the motor end stage of the decision process M1, the combined "lateralization signal" (i.e., regions from Figure 7A excluding M1), and the combined "searchlight signal" (i.e., all regions from Figure 7B). Critically, as predicted, the combined choice-specific signals, but not the M1 response, were significantly pushed towards the "yes"- choice (i.e., more positive in Figure 7D) for high compared to low TPR. The effect of TPR differed by cortical signal (2-way repeated measures ANOVA with factors signal type (3 levels) and TPR bin (2 levels); interaction: $F_{2,26} = 7.30$, $p=0.003$). Specifically, the difference of the choice-specific signals between low and high TPR was significantly larger for the combined lateralization signal and the combined searchlight signal than for M1 (combined lateralization signal vs. M1: $p=0.015$; combined searchlight signal vs. M1: $p=0.004$; permutation tests).

Because subjects' mean accuracy was about 74% correct, their choices were partially correlated with the physical stimulus (i.e., signal+noise vs. noise trials). Consequently, the choice-specific cortical responses were also (weakly) predictive of the stimulus (Figure 7–figure supplement 1H). To isolate variations in the amplitude of the choice-specific response that were independent of the stimulus, we removed (via linear regression) components explained by the stimulus and quantified the effect of TPR on the residual choice-specific cortical signals. Fitting the linear model to the combined choice-specific responses yielded highly significant TPR coefficients, for both the combined lateralization and combined searchlight signals (Figure 7E, middle and right panel). By contrast, the TPR-linked modulation was absent in the end stage region M1 (Figure 7E, left panel).

In sum, a number of fronto-parietal cortical regions exhibited signals that reliably encoded subjects' behavioral choice and were robustly modulated by phasic arousal, with a larger tendency towards the "yes"-choice under high TPR. This was true even when factoring out the effect of the sensory evidence (i.e. presence of the target signal).

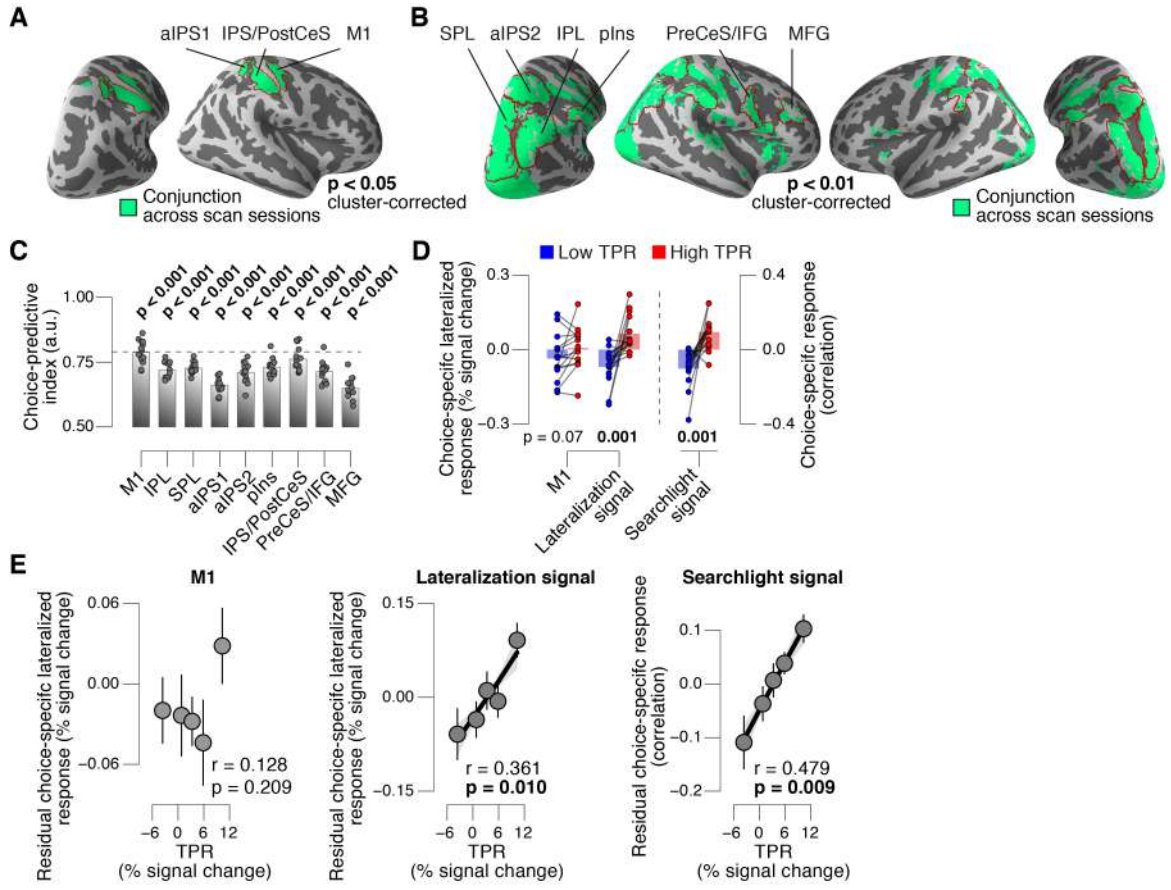


Figure 7: Phasic arousal predicts change of cortical decision signals. (A) Conjunction of session-wise maps of logistic regression coefficients of choice against fMRI lateralization (see Figure 7–figure supplement 1A for individual sessions). Tested against 0.5 at group level; red outlines, ROIs used for further analyses. **(B)** Conjunction of session-wise maps of searchlight choice classification precision scores (see Figure 7–figure supplement 1C for individual sessions). Tested against 0.5 at group level; red outlines, ROIs used for further analyses. **(C)** Choice-predictive indexes for choice-specific responses ("yes" vs. "no", irrespective of stimulus; see Materials and methods and Figure 7–figure supplement 1G). Dashed line, index for M1, which can be regarded as a reference given the measurement noise. Data points, individual subjects. **(D)** Choice-specific responses, obtained through mapping lateralization (M1 and the combined "lateralization signal", i.e., regions from Figure 7A excluding M1; see Materials and methods) and through searchlight classification (combined "searchlight signal", i.e., all regions from Figure 7B), for low and high TPR trials. Data points, individual subjects. **(E)** Correlation between TPR and M1 (left), or the combined "lateralization signal" (middle), or the combined "searchlight signal" (right) (5 bins). In all cases, the effect of the physical stimulus was removed (see Materials and methods). Shading or error bars, s.e.m. All panels: group average. (N = 14); stats, permutation test.

1.2.9 Task-evoked pupil response are predicted by responses in a network of brainstem centers

Finally, we aimed to identify brainstem regions whose task-evoked responses were (i) linked to the trial-to-trial fluctuations of TPR, and (ii) accounted for the trial-to-trial modulation of subjects' evidence accumulation bias, and the resulting tendency to choose "yes". Previous work from monkey physiology has implicated three brainstem nuclei in particular in the control of TPR: the locus coeruleus (LC), the inferior colliculus (IC), and the superior colliculus (SC), respectively [35, 36, 88]. Here, we exploited the wide coverage of our fMRI measurements to concurrently monitor responses across a wider brainstem

network, including a number of other nuclei implicated in central arousal: the dopaminergic substantia nigra (SN) and ventral tegmental area (VTA), as well as the (partly) cholinergic basal forebrain (BF). We further subdivided the BF region into the part including cell groups within the septum and the horizontal limb of the diagonal band (BF-sept) and the sublenticular part (BF-subl). BF-subl contains cholinergic neurons with widespread ascending projections [89], which are involved in the regulation of cortical arousal state [9, 11]. Our analysis approach minimized the effect of physiological noise on the brainstem fMRI responses, including removal of the fourth ventricle signal (see Materials and methods). We also verified that the fourth ventricle signal was unrelated to TPR (Figure 8–figure supplement 1D,E). The LC region of each subject was delineated through independent structural scans (Figure 8A, and Figure 8–figure supplement 1A; for details see Materials and methods).

The LC region exhibited a robust positive response on high TPR trials and a trend towards deactivation on low TPR trials (Figure 8B–D, and Figure 8–figure supplement 1C). The same pattern was evident for both signal+noise and noise trials separately (Figure 8C). The association to TPR was also highly significant in the most spatially specific definition of the LC region afforded by our measurements: evaluating only the two fMRI voxels with the largest probability of containing the individual LC region (Figure 8D, and see Materials and methods). Fluctuations of task-evoked fMRI responses measured in the LC were also robustly coupled to fluctuations in TPR amplitude at the single trial level (Figure 8F,H).

Similar to the LC region, we found a robust difference between low and high TPR conditions for fMRI responses in the SC and VTA regions (Figure 8E,F, and Figure 8–figure supplement 1B,C). Mapping the trial-to-trial correlations between TPR and brainstem fMRI responses at the single-voxel level yielded robust coupling to TPR in the LC, SC, VTA and as well as in BF-subl regions (Figure 8F).

As expected from the anatomical connectivity between brainstem centers [90–92], the trial-to-trial fluctuations of the task-evoked responses were significantly correlated among a number of these brainstem nuclei (Figure 8G). Removing components of the trial-to-trial fluctuations in TPR and fMRI responses shared with the other ROIs yielded significant residual (i.e., partial) correlations between TPR and responses in SC, LC region, VTA and BF-subl (Figure 8H). This indicates robust and unique contributions of these four nuclei to TPR.

Phasic brainstem responses during decision tasks might be driven by top-down signals from anterior cingulate cortex (ACC), which sends descending projections to the LC [6] and other brainstem nuclei. In line with this notion, trial-to-trial fluctuations of both LC responses and TPR were robustly correlated to trial-to-trial fluctuations of task-evoked responses of the ACC (Figure 8I).

1.2.10 Task-evoked responses in neuromodulatory centers, but not the colliculi, predict suppression of evidence accumulation bias

The task-evoked responses in the neuromodulatory nuclei, but not the colliculi, were tightly linked to the inferred decision computation and subjects' overt choice behavior. We computed the combined "neuromodulatory brainstem signal" as the linear combination of responses from LC, VTA, SN, and BF that

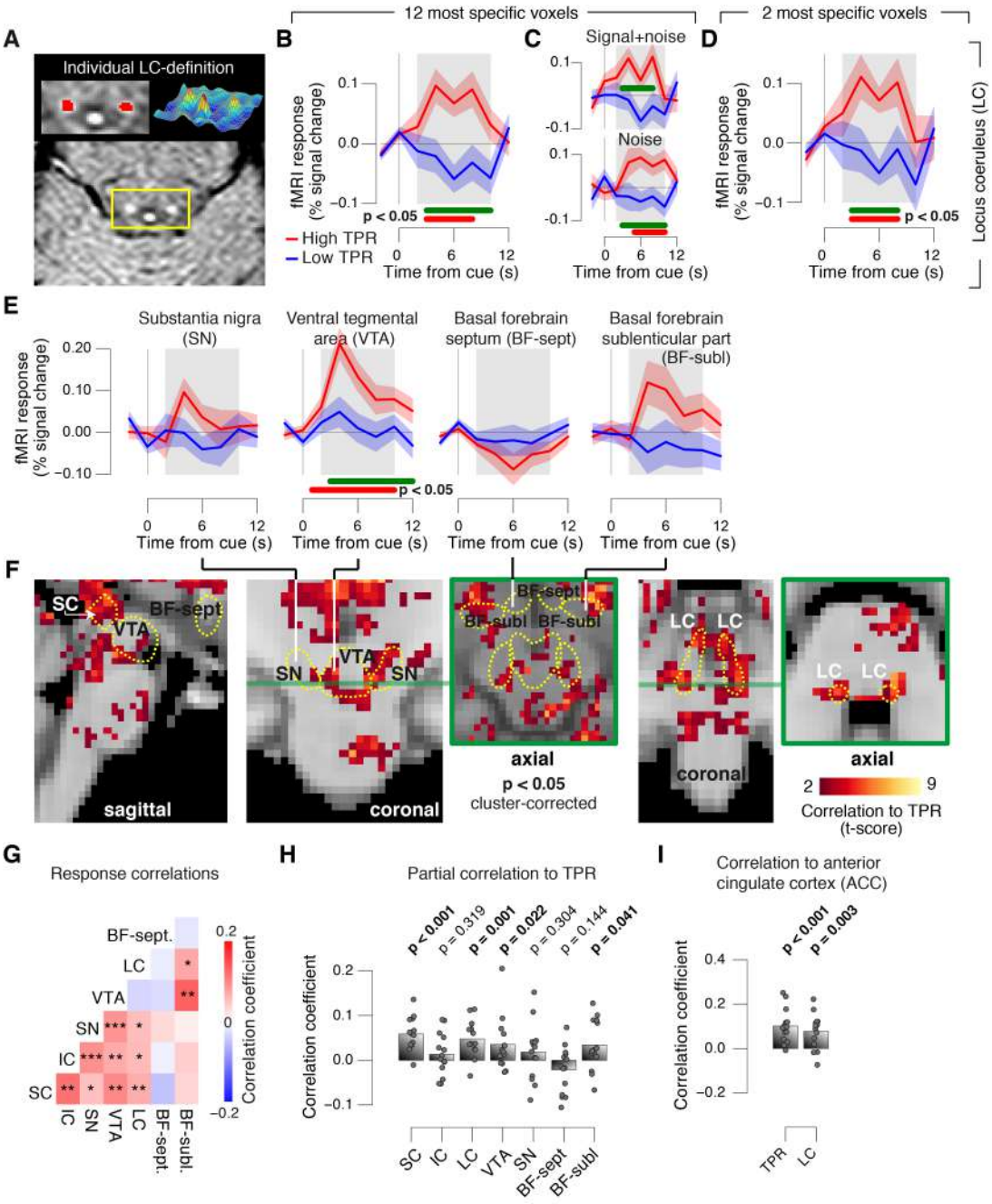


Figure 8: Pupil responses reflect responses of a network of brainstem nuclei. (A) Delineation of LC by structural scan. The LC corresponds to two hyper-intense spots; example subject (see Figure 8–figure supplement 1 for all subjects). Left inset, magnification of yellow box with LC ROI. Right inset, three-dimensional representation of signal intensity levels in yellow box. **(B)** Task-evoked LC responses for low and high TPR. Red bar, high TPR time course significantly different from zero; green bar, high TPR time course significantly different from low TPR time course ($p<0.05$; cluster-corrected). Grey box, time window for computing scalar response amplitudes. **(C)** As panel B, but split by signal+noise and noise trials. **(D)** As panel B, but for the 2 voxels with highest probability of containing the LC. **(E)** As panel B, but for SN, VTA, and two BF-ROIs. **(F)** Map of single-trial correlation between TPR and evoked fMRI responses (tested against 0 at group level). Yellow outlines, brainstem nuclei from probabilistic atlases. **(G)** Matrix of correlations between evoked brainstem fMRI responses. Stats corrected with false discovery rate (FDR). **(H)** Partial correlation of evoked fMRI responses and TPR. For each ROI, responses of all other ROIs were first removed via linear regression. **(I)** Correlation between fMRI responses in ACC and TPR and LC. All panels: group average (N = 14); shading, s.e.m.; data points, individual subjects; stats, permutation test.

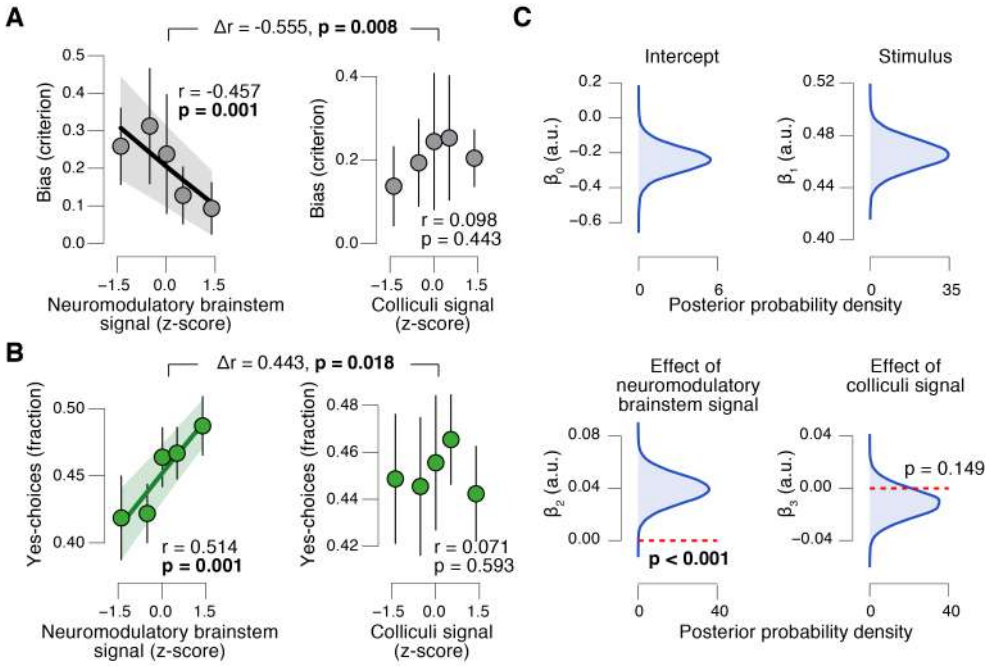


Figure 9: Brainstem neuromodulatory nuclei predict reduction of choice bias. **(A)** Correlation between decision bias (criterion) and the combined neuromodulatory brainstem signal (linear combination of responses in LC, SN, VTA, BF-sept, and BF-subl maximizing the correlation to TPR; see Materials and methods; left), and the combined colliculi signal (linear combination of responses in SC and IC maximizing the correlation to TPR; right) (5 bins). Stats, permutation test. **(B)** As panel A but for the correlation to fraction of "yes"-choices. **(C)** Group-level posterior probability densities for means of parameters in the DDM regression model, through which we assessed the trial-by-trial, linear relationship between single-trial drift and the combined neuromodulatory response or the combined colliculi response (see Materials and methods; see Figure 9–figure supplement 1 for the remaining parameters "starting point", "boundary separation" and "non-decision time"). All panels: group average (N = 14); shading or error bars, s.e.m.

maximized the correlation to TPR (Materials and methods; correlation coefficient across subjects, 0.146 (±0.014 s.e.m.)). The amplitude of this combined signal predicted a significant reduction in conservative decision bias (Figure 9A), and an increased tendency to choose "yes" (Figure 9B), but no change in sensitivity (Figure 9–figure supplement 1A). This pattern of effects was absent for the combined "colliculi signal" (Figure 9A,B), a linear combination of responses from SC and IC that maximized the correlation to TPR (correlation coefficient across subjects, 0.092 (±0.011 s.e.m.)). Further, the trial-to-trial variations in the strength of the combined neuromodulatory (but not colliculi) response robustly pushed the trial-to-trial drift towards the "yes"-boundary, in effect reducing the overall negative drift criterion (Figure 9D, see Materials and methods for details).

In sum, trial-to-trial fluctuations in TPR were predicted by fluctuations in the task-evoked responses of a network of brainstem regions, most notably the LC, VTA and SC. Despite the expected coupling between these and other brainstem regions (Figure 8G), TPR carried robust LC-, SC-, and (less strongly) VTA-specific components (Figure 8H). But only the responses of the neuromodulatory ROIs, not of the colliculi, accounted for the concomitant reduction of the bias in evidence accumulation and the resulting behavioral choice patterns. These results establish a tight link between phasic neuromodulator release and the dynamics of evidence accumulation.

1.3 Discussion

Intrinsic variability in the face of uncertain evidence is a pervasive feature of decision-making [1, 2, 4, 5]. Most current models of choice treat this intrinsic behavioral variability as a nuisance to be accounted for by additional "noise parameters" [18, 25]. Other theories have proposed that the behavioral variability may be due to hidden, but systematic, biases in the decision process [5, 67]. Here, we present evidence that helps reconcile these ideas. We found that a significant component of choice variability was explained by trial-to-trial variations in the amplitude of task-evoked, pupil-linked arousal responses. Specifically, pupil-linked arousal responses accounted for trial-to-trial variations in the bias of the evidence accumulation process as well as decision-related cortical population signals: under large phasic arousal conservative biases were reduced. The implication is that, without monitoring arousal responses, the associated, systematic variations in accumulation bias would appear as random trial-to-trial variability in the accumulation process (i.e., drift). Going further, we established that the dynamic bias suppression was explained by responses in a network of neuromodulatory brainstem systems controlling cortical arousal state. Taken together, our results are consistent with a scenario in which phasic neuromodulatory activity during decision-making optimizes choice behavior through a suppression of maladaptive biases in the evidence accumulation process.

1.3.1 Challenges and limitations of brainstem fMRI

Imaging the brainstem with fMRI is challenging [93–96] because this region is prone to physiological noise artifacts [95], and brainstem nuclei tend to be small relative to the spatial resolution of standard fMRI measurements. For example, although the adult human LC is an elongated structure of approximately 15 mm length along the rostro-caudal axis, its diameter is only a few millimeters, as assessed by high-resolution MRI (Figure 8A and Figure 8–figure supplement 1A) [97, 98]. Our study addressed these challenges by following the recommendations of Eckert and colleagues [99]: We (i) delineated the LC in each brain, based on individual (neuromelanin-sensitive) structural MRI scans; (ii) performed fMRI tailored to the anatomical layout of the LC while maximizing functional signal-to-noise ratio (SNR), by using an in-plane spatial resolution of 2x2 mm and 3 mm thick slices that were oriented perpendicular to the longitudinal extent of the LC; (iii) performed no spatial smoothing of these functional data; and (iv) rigorously removed measured cardiac and respiratory signal components, as well as residual fourth ventricle signal, which have been identified as a major source of uncertainty regarding previous fMRI work on the LC [93]. The resulting time-course of task-evoked fMRI responses exhibited the standard features of hemodynamic responses (Figure 8B–E), and correlations to pupil responses that are largely consistent with single-unit physiology in monkeys (see below). Taken together, the brainstem responses in Figures 8 and 9 likely reflect true neural signal from brainstem nuclei, rather than physiological noise. However, there is some inevitable uncertainty regarding the spatial specificity of our measurements. Due to the lower spatial resolution of fMRI images, the co-registration between functional and structural im-

ages, and the point spread of the hemodynamic response, each fMRI voxel is likely to sample activity from brain tissue neighboring the nuclei depicted as the regions of interest (e.g., LC). Consequently, we do not conclude that the LC responses in Figure 8B–D reflect the activity of noradrenergic neurons only; such a conclusion would require single-unit measurements. The focus of our conclusions instead lies on the distribution of pupil and behavioral correlations across different brainstem structures, which provides an important complement to targeted single-unit measurements.

1.3.2 Brainstem correlates of pupil dilation and decision bias

Despite the above-mentioned limitations, the overall distribution of pupil-linked brainstem responses shown in Figure 8F meaningfully follows the outlines of key candidate structures, in a fashion that is largely consistent with monkey physiology, and a previous human study on fMRI correlates of fluctuations in baseline pupil diameter [34]. Our approach also identified so-far unknown effects. Previous monkey physiology has established significant coupling of pupil responses to responses of the LC, SC, and IC [35, 36, 88], but not yet for dopaminergic and cholinergic structures (i.e., SN, VTA, and BF). The ability to monitor all of the above brainstem regions at once enabled quantification of their trial-to-trial correlation structure, and hence isolating the contributions that were unique to each region. This revealed that (i) many brainstem nuclei co-fluctuated during the decision task and that (ii) not only the LC and SC, but also the VTA and sublenticular part of the BF each made robust and specific contributions to task-evoked pupil dilations, over and above those shared with other brainstem centers (Figure 8H). Thus, the noradrenergic, cholinergic and dopaminergic systems are all phasically, and to some extent independently, recruited during challenging decision tasks, and jointly shape the concomitant changes in arousal state. Our findings provide a basis for a more comprehensive neurophysiological interpretation of results from cognitive pupillometry studies in humans.

Most importantly, we also established that only a subset of those brainstem nuclei exhibiting robust correlations with pupil responses were also predictive of the trial-by-trial suppression in decision bias. The latter effect was solely accounted for by responses in the (noradrenergic, dopaminergic, and cholinergic) neuromodulatory nuclei with diffuse projections to cortex, but not by responses in the superior or inferior colliculi. This indicates that the phasic release of neuromodulators in the brain, possibly a combination of different neuromodulators, is key for behavioral correlates of phasic arousal identified here.

1.3.3 Phasic versus tonic arousal effects

A number of recent studies have characterized the relationship between tonic arousal levels (measured through baseline pupil diameter) and cortical state [76, 78, 79, 100]. Other studies have characterized the relationship between tonic arousal levels and behavioral performance [76, 101]. The comparison between this previous work and ours points to possible differences between the functional correlates of tonic arousal levels and phasic, task-evoked changes in arousal. We found that phasic, task-evoked arousal re-

sponses were primarily linked to decision bias, at both, the algorithmic and cortical levels. By contrast, the above studies of tonic arousal levels have revealed effects on the quality of sensory cortical responses and behavioral sensitivity to sensory evidence [11]. While we also found some evidence for non-monotonic (inverted U-shape) relationships between phasic arousal and sensitivity, the dominant and most consistent link was a monotonic and approximately linear relationship between phasic arousal and decision bias. Candidate factors accounting for these apparent differences between the functional correlates of phasic and tonic arousal might be the dynamics of the underlying neuromodulatory effects on cortical circuits, or the different combination of neuromodulatory systems involved. It will be instrumental to track TPR-linked changes in brainstem and cortical state in real time in future work.

1.3.4 Post-decisional versus intra-decisional drive of phasic arousal

One account holds that the phasic arousal signals (specifically, phasic responses of the noradrenergic LC) are triggered by the bound-crossing in one of the cortical accumulator circuits; the resulting transient and cortex-wide neuromodulator release then facilitates the translation of the choice into a motor act [6]. An alternative idea [30], supported by indirect evidence [40, 43], is that arousal systems are already recruited before the bound-crossing, throughout the evidence accumulation process. In line with the latter notion, we found that task-evoked pupil responses are driven most strongly by a sustained central input throughout decision formation, not only after commitment to a choice. This finding has potentially important implications for the functional role of phasic arousal in decision processing. The finding indicates that at least one of the brainstem nuclei linked to pupil responses was, likewise, activated in a sustained fashion throughout decision formation. The resulting neuromodulatory transients might alter the state of brain regions involved in decision computations as the decision unfolds, provided that the accumulation operates on timescales of seconds or longer. Because the tasks used in previous animal physiology studies of task-related LC responses involved much faster decision processes than the one studied here (reaction times of about 0.5 s vs. 2 s, respectively), it remains unknown whether the more sustained, task-evoked responses also occur in noradrenergic neurons (but see [36]). Sustained responses encoding reward uncertainty have been observed in dopaminergic neurons in the VTA [102], one of the structures whose task-evoked responses predicted pupil responses. Future electrophysiological studies should determine the time course of task-related activation in the different nuclei of the brain's arousal network during sensory-motor decisions involving protracted evidence accumulation [103].

1.3.5 How do diffuse neuromodulatory signals translate into specific effects on choice behavior?

One notable aspect of our findings is that the functional correlates of neuromodulatory responses were specific for a particular choice option (see Figure 9). Also in the context of learning, the interplay between pupil-linked arousal and competitive cortical circuitry has been found to translate into specific effects on cognition and behavior [104].

A scenario consistent with our results is that phasic neuromodulator release alters the relative strength of information flow between cortical processing stages, suppressing "top-down" relative to "bottom-up" signals [57–61]. In perceptual decisions like the ones studied here, early sensory cortices provide bottom-up sensory likelihood signals, while top-down signals might encode prior beliefs [57, 62]. Thus, through a relative suppression of "top-down" signal flow, phasic arousal might reduce the weight of the prior (reflecting subjects' intrinsic bias) relative to the likelihood. Specifically, in our yes-no task, the prior may have been a conservative bias for choosing "no". Reducing its weight would reduce this bias. Such an increase in the relative weight of bottom-up signals might be implemented by synaptic gain modulation through neuromodulators. This gain modulation, in turn, might depend on the precision (inverse of uncertainty) of incoming sensory data [57, 105].

The above scenarios postulate a uni-directional effect of neuromodulatory transients on cortical decision computations. However, this interaction may also be bi-directional, with trial-to-trial fluctuations of cortical decision signals driving fluctuations of phasic arousal responses [6, 30]. Specifically, phasic LC responses may be driven by specific cortical regions (e.g., the ACC), which compute the ratio of the posterior probability of target presence over the (estimated) prior probability of target occurrence [30]. The resulting phasic norepinephrine release across cortex might reset cortical networks [31] and interrupt the (default) state encoding the prior [30]. In a yes-no task such as ours, a tendency towards the "no"-option may correspond to the default state for conservative subjects, and a phasic arousal signal is generated when decision-related neural activity ramps towards "yes", facilitating the transition of the entire cortical system towards that non-default state.

1.3.6 Conclusion

Our findings establish that phasic task-evoked pupil responses during the formation of sensory-motor decisions reflect responses of a network of neuromodulatory brainstem centers including the noradrenergic LC. Phasic, pupil-linked arousal alters choice-encoding population signals in parietal and prefrontal association cortices. Phasic arousal in general, and neuromodulatory brainstem responses in particular, explain a dynamic reduction in decision-makers' bias towards one particular choice. The resulting trial-to-trial variability of decision bias accounts for a significant component of the intrinsic behavioral variability: when decisions are made in the face of uncertainty, tracking phasic arousal signals may be just as important for predicting choice behavior as tracking the objective evidence gathered from the outside world.

1.4 Materials and Methods

1.4.1 Subjects

We report analyses of four independent data sets, from behavioral tasks described in the subsequent section. All subjects had normal or corrected-to-normal vision and gave written informed consent. Subjects

received e15 per hour (all visual tasks) or research credit (auditory task) for their participation. The ethics committee of the Psychology Department of the University of Amsterdam approved the experiments.

Fifteen healthy subjects (5 females; age range, 22–35 y) participated in the main experiment of this study, entailing concurrent pupillometry and brainstem as well as cortical fMRI recordings. Here, each subject participated in several fMRI sessions: one to define retinotopically organized visual cortical areas (75 min) and two sessions (three for one subject) for the main experiment (about 2 hr per session). Three subjects were authors, and the remaining 12 subjects were naive to the purpose of the study. The results were unchanged when excluding the three authors (see Author response, online) and the one subject who performed three sessions (and more trials; see section *Behavioral tasks*) of the main experiment. One (male) subject was excluded from the analyses because the stimulus software did not receive the triggers from the MRI scanner in two sessions (the age range remained the same).

We also re-analyzed the 21 subjects from an existing behavioral data set, for which we had previously published different analyses [40] (Figure 2, Figure 2–figure supplement 1, Figure 4 and Figure 4–figure supplement 1). In that experiment, 23 subjects had performed a yes-no visual contrast detection task with trial structure analogous to that of the fMRI experiment, enabling joint fitting of the drift diffusion model to both data sets using a hierarchical Bayesian procedure (see below). To this end, we excluded the two subjects from the [40] data set who had also participated in the current fMRI experiment, keeping the two samples independent.

Finally, 24 subjects (20 females; age range, 19–23 y) performed an auditory tone-in-noise detection task (Figure 3A,B), and 15 subjects (six females; age range, 23–37 y) performed a visual random dot motion discrimination task (Figure 3C,D).

1.4.2 Sample sizes

The sample sizes were determined based on a number of criteria: (i) the assessment of the behavioral correlates of TPR obtained in a previous study [40]; (ii) the need to obtain as many trials as possible from each individual (necessary for detailed modeling of choice behavior as a function of TPR within subjects at a first level, before second-level statistics; Figures 2–4); and, in the case of fMRI, (iii) the need to obtain detailed retinotopic maps per individual from a separate scanning session (Figure 5), as well as robust maps of choice-specific activity by means of conjunction across two sessions of the main experiment (Figure 7, and Figure 7–figure supplement 1). Taken together, these criteria prioritized obtaining a large amount of data (and experimental sessions) from each participant, which was traded off against the total number of participants.

1.4.3 Behavioral tasks

Main task: Visual contrast detection (yes-no) Each trial began with the central fixation dot turning green and consisted of three consecutive intervals (Figure 1A): (i) the baseline interval (2 s; containing only noise); (ii) the decision interval, the start of which was signaled by the occurrence of a tone

(200 ms duration) and which was terminated by the subject's response (or after a maximum duration of 3.5 s); (iii) the inter-trial interval (ITI), which consisted of a dark grey fixation dot on an otherwise blank screen and was uniformly distributed between 4 and 12 s.

A dynamic noise pattern (refresh rate: 60 Hz) was presented throughout the trial. The luminance across all pixels was kept constant. This pedestal noise pattern had 10% contrast and was refreshed on each frame. On one half of trials ("signal+noise" trials), a sinusoidal grating (five cycles per degree) was superimposed on the visual noise for the entire decision interval, from the onset of the auditory cue to the subject's motor response (Figure 1A). The other half of trials ("noise" trials) contained no target signal during the decision interval. Signal presence was randomly selected on each trial, under the constraint that it would occur on 50% of the trials within each block of 40 trials. All stimuli were presented in a Gaussian annulus, with an average distance (±SD) to fixation of 1.8 (0.6) degrees (Figure 1A). In different blocks, the target grating, if present, was tilted 45°(clockwise, CW) or 135°(counter-clockwise, CCW). To minimize uncertainty, subjects were informed about the orientation of the target before each block, by means of a full-contrast presentation of the target signal.

Subjects 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" –> press right key; "no" –> press left key) was counterbalanced across subjects. At the end of each block of 40 trials subjects were informed about their performance.

Throughout the main experiment, the contrast of the target signal was fixed at a level that yielded about 75% correct choices. Each subject's individual threshold contrast was determined before the main experiment in the MRI scanner (during anatomical scans), using an adaptive staircase procedure (Quest). Here, we used a two-interval forced choice variant of the contrast detection task (one interval: signal+noise, the other: noise). The corresponding threshold contrasts yielded a mean accuracy of 73.79% correct (±1.32% s.e.m.) in the yes-no visual contrast detection task during fMRI. Subjects performed between 10 and 12 blocks (distributed over two scanning sessions), yielding a total of 400–480 trials per subject. One subject performed a total of 16 blocks (distributed over three scanning sessions), yielding a total of 640 trials.

Stimuli were back-projected on a transparent screen using a gamma-corrected LCD projector with a spatial resolution of 1920 × 1200 pixels, run at a vertical refresh rate of 60 Hz. Subjects were supine in the MRI scanner and viewed the screen from 120 cm via a mirror attached to the head coil. To minimize any effect of light on pupil diameter, the overall luminance of the screen was held constant throughout the experiment.

Auditory tone detection task (yes-no) The trial structure was identical to that of the visual contrast detection task described above, but without noise during the pre-decision interval baseline, and with shorter ITIs (uniformly distributed between 1 and 2 s). The decision interval consisted of an auditory noise stimulus (as in [76]), or a pure sine wave (2 KHz) superimposed onto the noise (50% of trials each).

These stimuli were presented from the onset of the auditory stimulus to the subject's motor response. Threshold volumes (in dB), determined beforehand via an adaptive staircase procedure, yielded a mean accuracy of 73.41% correct (±0.90% s.e.m.). Auditory stimuli were presented using an IMG Stageline MD-5000DR over-ear headphone, suppressing ambient noise. Subjects performed between 11 and 13 blocks in the behavioral lab (distributed over two measurement sessions; same set-up as in [40]), yielding a total of 1320–1560 trials per subject.

Visual motion discrimination task (two-alternative forced choice) The trial structure was identical to that of the visual contrast detection task described above, but with fixed stimulus duration (750 ms; interrogation protocol) and visual feedback at the end of each trial (green/red rectangle at fixation to signal correct/error). An auditory white noise stimulus (250 ms) was played on 50% of the trials. The mapping between perceptual choice and button press (e. g., "up" –> press right key; "down" –> press left key) varied from session to session, counterbalanced within subjects.

Random dot motion stimuli (refresh rate: 120 Hz) were presented throughout the experiment in a central annulus (outer diameter 16.8°, inner diameter of 2.4°) around the central fixation rectangle (0.45°length). The annulus contained 524 dots all within one hemifield (half circle). Hemifield presentation (left or right) changed across blocks and was counterbalanced across subjects and sessions. Dots were 0.15°in diameter and white, presented on a grey background. Dots were divided into "signal" and "noise" dots, the proportion of which defined the motion coherence level. Outside the stimulus interval, motion coherence was fixed to zero (pure noise). During the stimulus interval, signal dots moved at 7.5°/s in one of two directions (up or down), were randomly selected on each frame, had a lifetime of 10 frames, and were re-plotted in random locations thereafter (reappearing on the other side when their motion extended outside of the annulus). Noise dots were randomly assigned to locations within the annulus on each frame. Independent motion sequences (n = 3) were interleaved to prevent tracking of individual dots [106]. Threshold motion coherence, determined beforehand via an adaptive staircase procedure, levels yielded a mean accuracy of 88.86% correct (±1.49% s.e.m.) and 71.01% correct (±1.35% s.e.m.), respectively. We collapsed across these two peri-threshold levels, to maximize statistical power, but we verified that the effect found was analogous for both difficulty levels.

Subjects performed between 23 and 24 blocks (distributed over four sessions; same set-up as in the main task, except that stimuli were presented on a 31.55' MRI compatible LCD display with a spatial resolution of 1920 × 1080 pixels) yielding a total of 575–600 trials per subject. One subject performed a total of 18 blocks (distributed over three scanning sessions), yielding a total of 450 trials. Here, we analyzed only the 50% of trials without the auditory white noise stimulus (see above), which are comparable to the trials from the other tasks. Data from one measurement session of one subject was excluded from the analyses because of poor eye-tracker data quality.

1.4.4 Magnetic resonance imaging data acquisition

For the main experiment, MRI data were acquired on a 3T Philips Achieva XT MRI scanner using a 32-channel head coil in two types of sessions: retinotopic mapping sessions (for defining the borders of visual cortical areas V1-V3, see section *Definition of regions of interest*) and main experimental sessions. In all sessions, cardiac cycle was monitored with a pulse oximeter attached to the left index finger, and respiratory activity was recorded with a chest belt, for physiological noise removal. Both physiological signals were recorded at a sampling rate of 496 Hz.

During the main experimental sessions, EPI images were acquired in 35 slices (thickness: 3.0 mm, no gaps) oriented perpendicular to the floor of the fourth ventricle (i.e., perpendicular to the longitudinal extent of the locus coeruleus [97], with the following parameters: TR = 2 s, TE = 27.62 ms, flip angle = 76.1°, SENSE acceleration factor = 3.0. Images were acquired at an inplane resolution of 2.0×2.06 mm and were reconstructed at a resolution of 1.79×1.79 mm. A structural T1 scan was acquired with an MPRAGE sequence for anatomical co-registration and cortical surface reconstruction (voxel size: $1 \times 1 \times 1$ mm, TR = 8.2 ms, TE = 3.73 ms, flip angle = 8°). An additional structural scan was acquired with a T2-weighted sequence and higher resolution than the EPI scans ($1 \times 1 \times 1.5$ mm, TR = 5114 ms, TE = 12.5 ms, flip angle = 90°) to facilitate co-registration of EPI images and the high-resolution structural T1 scan. Two turbo spin echo (TSE) neuromelanin-sensitive structural scans were acquired for delineation of the LC [97, 98, 107], again oriented perpendicular to the floor of the fourth ventricle. The first (partial field-of-view) TSE scan was obtained with the following parameters: 20 slices (1.5 mm, no gaps), in-plane resolution: 0.7×0.88 (reconstructed at: 0.35×0.35 mm), TR = 500 ms, TE = 10 ms, flip angle = 90°, covering the brainstem only. The second (whole-brain) TSE scan was obtained with the following parameters: 35 slices (3 mm, no gaps), in-plane resolution: 1.96×2.0 (reconstructed at: 0.47×0.47 mm), TR = 500 ms, TE = 10 ms, flip angle = 90°. Finally, field maps were acquired using two separate acquisitions (voxel size: $2 \times 2 \times 2$ mm3, TR = 11 ms, TE1 = 3.0, TE2 = 3.5, ms, flip angle = 8°).

During retinotopy sessions, EPI scans were acquired in 29 slices (thickness: 2.5 mm, with 0.25 mm slice gaps) with the following parameters: in-plane resolution: 2.5×2.58 mm (reconstructed at 2.5×2.5 mm), TR = 1.5 s, TE = 27.62 ms, flip angle = 70°, SENSE acceleration factor = 3.0. An additional structural scan was acquired with a T2-weighted sequence and higher resolution than the EPI scans ($1.25 \times 1.25 \times 1.25$ mm with 0.12 mm slice gaps, TR = 8390 ms, TE = 100 ms, flip angle = 90°) to facilitate co-registration of EPI images and the high-resolution structural T1 scan used for cortical surface reconstruction (see above).

1.4.5 Eye data acquisition

Concurrently with the fMRI recordings, the left eye's pupil was tracked (via the mirror attached to the head coil) 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). The MRI-compatible (non-ferromagnetic) eye tracker was placed outside the scanner bore, and it was calibrated once at the start of each scanning session. The purely behavioral experiments were conducted in a psychophysics laboratory. Here, the left

eye's pupil was also tracked at 1000 Hz with an average spatial resolution of 15 to 30 min arc, using the same EyeLink 1000 system (SR Research, Osgoode, Ontario, Canada).

1.4.6 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 software. We applied to each pupil recording (i) linear interpolation of values measured just before and after each identified blink (interpolation time window, from 150 ms before until 150 ms after blink), (ii) band-pass filtering (third-order Butterworth, passband: 0.01–6 Hz), (iii) 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 [108], and (iv) conversion to units of modulation (percent signal change) around the mean of the pupil time series from each block. Filtering (specifically the lower cutoff at 0.01 Hz) was performed on both fMRI and pupil time series, ensuring equal treatment of both signals to be correlated.

Quantification of task-evoked pupillary responses (TPR) We computed task-evoked pupillary response (TPR) amplitude measures for each trial as the mean of the pupil diameter modulation values in the window –1 s to 1.5 s from choice (same time window as in [40]), minus the mean baseline pupil value during the 0.5 s before the cue (i.e., decision interval onset) (Figure 1B). The sluggish hemodynamic system and, to a lesser extent, the peripheral pupil apparatus [41, 109] act as temporal low-pass filters. As a result, trial-to-trial variations in reaction time (RT; and thus, the duration of the task-related sustained activity, see Figure 1D, and Figure 1–figure supplement 1A–C) can induce trial-to-trial variations of the fMRI and TPR amplitudes, without changes in the amplitude of the underlying neural responses. Indeed, there was a robust relationship between reaction time and TPR (Figure 1–figure supplement 1D). Therefore, to specifically isolate trial-to-trial variations of underlying response amplitudes, variations due to RT were removed via linear regression from both the TPR and fMRI responses. This was done for all analyses reported in the main text. Hereby, RT was defined as the time from decision onset (cued by tone) until the button press. To establish the robustness of the effects reported here we verified that all pupil-linked behavioral results were evident also without removing RT-related components (Figure 2–figure supplement 1, and Figure 4–figure supplement 2).

In the majority of analyses, trials were sorted by TPR amplitude and collapsed into three bins containing the lowest and highest 40% (which were used for analyses), as well as the intermediate 20% of TPR amplitudes (Figure 1B,C). This achieved a trade-off between maximizing both (i) trial counts in the high and low TPR bins and (ii) the disparity between the TPR amplitudes for both bins. In other analyses, we used five equally populated bins of single-trial TPR amplitudes (Figure 1–figure supplement 1, Figure 2, Figure 2–figure supplement 1, Figure 3 and Figure 7). In Figure 3D, we used four bins because of the lower trial count per subject in this data set after excluding the trials with the auditory white noise

manipulation (see section *Behavioral tasks*)

General linear modeling of TPR We used a general linear modeling approach described in detail in a previous paper [40] to estimate the relative contribution of three different putative temporal input components to the peripheral apparatus controlling pupil motility [110]. We (i) cut the pupil time series into single-trial epochs ranging from 1 s before cue to 5 s after cue, (ii) baseline-corrected each epoch by subtracting the mean baseline pupil value during the 0.5 s before cue, and (iii) concatenated these baseline-corrected epochs into a new time series excluding large parts of the inter-trial intervals. The GLM was then fit to this new, cleaned-up time series. The GLM consisted of the following transient events (Figure 1D, and Figure 1–figure supplement 1A–C): cue (onset of decision interval) and the choice. The transient corresponding to choice was placed at 0.24 s before button press, adopted from a report quantifying the interval between phasic LC activity and behavioral response in a forced-choice task in monkey [47]. The GLM also consisted of a sustained component in between the two transient events, which was modeled as a boxcar function. We normalized the boxcar regressor by dividing the height of the boxcar by the number of samples in that particular interval, such that this regressor had the same norm as the transient regressors. Thus, estimated beta weights were comparable between both sets of regressors. Each regressor was then convolved with a canonical pupil impulse response function (parameters taken from [41]), and multiple regression yielded the best-fitting beta weights for each regressor type (i.e., temporal component of the pupil response), separately for each subject.

1.4.7 Analysis and modeling of choice behavior

The first trial from each block and trials in which subjects failed to respond within the time limit of 3.5 s (see section *Stimuli, task and procedure*) were excluded from all analyses. RT was defined as the time from decision interval onset (cued by tone) until the button press. In a model-free analysis, we computed the fraction of "yes"-choices separately for two TPR bins (Figure 2, Figure 2–figure supplement 1, and Figure 3. To ensure that each TPR bin consisted of the same number of signal+noise and noise trials, we (i) sorted all trials of a subject into of four "cells" defined by the factors TPR (low and high) and stimulus (signal+noise and noise), (ii) determined the lowest trial count across the four cells, (iii) randomly sampled the same number of trials (without replacement) from the remaining cells, and (iv) computed the fraction of "yes"-choices separately for the two TPR bins. We then repeated this procedure 1000 times and averaged the results across all repetitions. The fraction of non-preferred choices (Figure 3C) was computed in the same way, with the exception that the non-preferred choice was defined as the choice opposite to the subject's overall bias (towards up or down) calculated across all trials (i.e., irrespective of TPR). We then modeled the effects of phasic arousal (as indexed by TPR) on choice behavior using two approaches, which yielded converging results.

Signal-detection theoretic (SDT) modeling In a first approach, we computed the SDT-metrics d' and criterion [24] separately for multiple (two or five) bins of TPR (Figure 2, Figure 2–figure supplement 1, and Figure 3) or combined brainstem response (Figure 9, and Figure 9–figure supplement 1). We estimated d' as the difference between z-scores of hit- and false-alarm rates. We estimated criterion by averaging the z-scores of hit- and false-alarm rates and multiplying the result by –1.

We used sequential polynomial regression analysis [111] to quantify the dependence of d' and criterion on TPR. This procedure allowed us to systematically test whether TPR predominantly exhibited no (zero-order polynomial), a monotonic (first-order polynomial), or a non-monotonic (second-order polynomial) effect on the SDT metrics. The SDT metric y was modeled as a linear combination of polynomial basis functions of 5 TPR bins:

$$y \sim \beta_0 + \beta_1 TPR^1 + \beta_2 TPR^2 \quad (1)$$

with β as polynomial coefficients. The corresponding regressors were orthogonalized, and each model was sequentially tested in a serial hierarchical analysis, based on F-statistics. This analysis was performed at the group level, 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 TPR) up to the second-order (quadratic, non-monotonic). If the first-order model was significantly better than the zero-order model at the group level, we fitted a linear model and tested the corresponding linear correlation coefficients across the group. This was true for SDT criterion in all cases, except Figure 3D, in which the linear expansion was only marginally significant; however, the linear correlation was highly significant when tested across the group. If the second-order model was significantly better than the first-order model at the group level, we fitted a quadratic model between TPR and behavior for each subject and tested the second-order coefficients across the group. This was true for SDT d' in some of the cases (Figures 2E and 3B).

Having established a robust first-order (monotonic) relationship between TPR and SDT criterion, we then characterized the timing of this effect by means of a sliding window (linear) correlation analysis over the interval from 1 s before cue to 3 s after response (window length: 250 ms, step size: 25 ms). We computed separate, baseline-corrected TPR values (see section *Quantification of task-evoked pupillary responses*) for each position of the window. Per time window, we then sorted trials by the TPR-values into five bins, and correlated these values with criterion estimates for the corresponding bins. This yielded time courses of the correlation between TPR and criterion.

Drift diffusion modeling In the second approach, we fitted the drift diffusion model [25] to RT distributions for "yes"- and "no"-choices, separately for low and high TPR trials (Figure 4, Figure 4–figure supplement 2). We fitted the model using the hierarchical Bayesian implementation of the HDDM toolbox [77] (version 0.6). The group distribution constrains individual subject parameter estimates, with a stronger influence when its variance is estimated to be small (for details of the procedure, see [77]).

Fitting the model to RT distributions for "yes"- and "no"- choices (termed "stimulus coding" in [77]), as opposed to the more common fits of correct and incorrect choice RTs (termed "accuracy coding" in [77]), was essential for estimating parameters that could have induced biases in subjects' behavior.

We fit the model to the behavioral data from a total of 35 subjects: 14 subjects from the current fMRI study, and 21 subjects from a previous study employing an analogous contrast detection task ([40]; see section *Subjects*). Doing so improved the estimation of the group-level distribution over parameters, which was used to constrain the individual subject parameter estimates. Specifically, the pooling was required to jointly fit the parameters starting point and drift criterion, both of which can account for changes in decision bias, but via different mechanisms, and are distinguishable through their distinct effects on the shape of the RT distribution (Figure 4–figure supplement 1). In our fits, we allowed the following parameters to vary between low and high TPR: (i) the mean drift rate across trials; (ii) the drift criterion (an evidence-independent constant added to the drift toward one or the other bound); (iii) the separation between both decision bounds (i.e., response caution); (iv) the starting point of the accumulation process; (v) the non-decision time (sum of the latencies for sensory encoding and motor execution of the choice). Drift rate variability is an additional parameter that was found to improve fits to empirical RT distributions [25]. However, this parameter is prone to fit error [112] and we had no a priori hypothesis about its relationship to phasic arousal. We therefore fit drift rate variability to all data (i.e., regardless of TPR), to maximize robustness of our fits.

To test the robustness of the significance of the TPR-dependent effect on drift criterion, we re-fitted the model, but now fixing drift criterion with TPR, while still allowing all other of the above parameters to vary with TPR. Using deviance information criterion [113] for model selection, we compared whether the added complexity of our original model was justified to account for the data. This is a common metric for comparing hierarchical models, for which a unique "likelihood" is not defined, and the effective number of degrees of freedom is often unclear [113].

Finally, we used the HDDM toolbox [77] to assess the trial-by-trial, linear relationship between the combined neuromodulatory brainstem responses, the combined colliculi responses, and the drift (Figure 9C, Figure 9–figure supplement 1). We fitted a variant of the above-described model (Figure 4A) to RT distributions for "yes"- and "no"-choices from the 14 subjects from the current fMRI study, but now modeling the drift on each trial as the following linear combination:

$$v \sim \beta_0 + \beta_1 S + + \beta_2 M + \beta_3 C \tag{2}$$

where v was the single-trial drift, S was a binary vector describing the stimulus identity (1, signal+noise; –1, noise), M was a vector of the single-trial combined neuromodulatory brainstem response (see section *Quantification of task-evoked fMRI responses and correlation with TPR*), C was a vector of the single-trial combined colliculi response. The fit parameters quantified how the drift on single trials was affected by the overall drift bias (i.e., mean drift criterion across trials, regardless of brainstem response, β_0), the overall drift rate (i.e., mean stimulus-dependent drift across trials, β_1), and the neural responses of the combined

neuromodulatory nuclei or colliculi (β_2 and β_3), respectively. The parameters starting point, boundary separation, and non-decision time were also included in the model, but not as a function of either of the combined brainstem responses.

1.4.8 Analysis of MRI data

MRI data were analyzed using custom-made software written in Python. A number of processing steps relied on FSL [114] and FreeSurfer [115, 116].

Preprocessing T1-weighted anatomical scans acquired at the beginning of each scanning session were automatically segmented and inflated for visualization using FreeSurfer [115, 116]. We then applied to the EPI scans (i) removal of non-brain tissue (brain extraction) using the BET tool in FSL, (ii) unwarping using a B0 field map and FUGUE (FMRIB's Utility for Geometrically Unwarping EPI's), (iii) image realignment to compensate for small head movements [117] (for improved precision we used as the target volume the high-resolution T2-weighted anatomical scan; the resulting up-sampled EPI scans were down-sampled back to their original resolution), (iv) high-pass filtering to correct for baseline drifts in the signal (Gaussian-weighted least-squares straight line fitting, with window size = 50 samples), and (v) conversion to units of modulation (percent signal change) around the mean fMRI series. We concatenated all EPI volumes preprocessed in that way across blocks within one scanning session. We applied physiological noise correction using FSL PNM [95], an extended version of RETROICOR [118], whereby cardiac and respiratory phases were assigned, separately for each slice, to each volume in the concatenated EPI image time series.

Our complete physiological noise regression model included 34 physiological noise regressors [95]: 4th order harmonics to capture the cardiac cycle, 4th order harmonics to capture the respiratory cycle, 2nd order harmonics to capture the interaction between the cardiac and respiratory cycles, 2nd order harmonics to capture the interaction between the respiratory and cardiac cycles, one regressor to capture heart rate, and one regressor to capture respiration volume. These 34 noise predictors, plus two for fMRI responses to eye-blinks and saccades (obtained by convolving blink and saccade events with a canonical hemodynamic impulse response function), were regressed against the time series of EPI volumes using multiple linear regression. The residual (i.e., noise-corrected) time series were used for all further analyses.

Subsequent analyses proceeded along two separate pipelines: (i) in functional space, by extracting responses from several, specifically delineated regions of interest (ROIs); or (ii) in anatomical space, as voxel-wise functional maps. The procedures for the delineation of ROIs are described in the section *Definition of regions of interest (ROIs)* below. For the computation of group average TPR-brainstem correlation map (Figure 8F) and of the cortex-wide functional maps (Figure 6, Figure 7A,B, and Figure 7–figure supplement 1A–D), the time series of EPI volumes were first transformed to MNI space (affine transformation with 12 degrees of freedom and sinc interpolation with FSL FLIRT). For the computation of cortex-wide functional maps, we additionally applied spatial smoothing to all EPI volumes in MNI space

using isotropic Gaussian filter kernels; the full width at half maximum (FWHM) of the kernels was 8 mm for all maps except for those of searchlight precision scores (Figure 7B, and Figure 7–figure supplement 1C,D). All other analyses reported in this paper were performed without spatial smoothing.

Quantification of task-evoked fMRI responses and correlation with TPR The slow event-related design (mean ITI: 8 s) enabled us to quantify task-evoked fMRI responses for each trial as the difference between fMRI measurements during the trial (starting from 2 s from cue) and the mean fMRI response during the pre-decision baseline interval (–2 s to 2 s from cue). Scalar fMRI response amplitudes were computed by collapsing response values across the interval 2 s to 12 s from cue. Trial-to-trial variations of each voxel's task-evoked fMRI response amplitude due to RT were removed (via linear regression), to isolate fMRI response variations that were due to variations in the amplitude of the underlying neural responses (see section *Analysis of task-evoked pupil responses*). For the analysis of evoked fMRI responses in the brainstem (Figure 8, and Figure 8–figure supplement 1), we additionally removed (via linear regression) signal fluctuations from the fourth ventricle (delineated based on the TSE scan by averaging across all voxels covering the ventricle) from the time series from each brainstem ROI or voxel, before computing task-evoked responses.

We computed maps of task-evoked fMRI responses as the voxel-wise difference between fMRI response during the decision interval and during baseline. The resulting maps were computed separately for each subject and then tested against 0 across the group (see section *Statistical comparisons*). We computed the following statistical maps: (i) overall task-evoked response (Figure 6A), (ii) difference between high and low TPR trials (Figure 6B), (iii) difference between signal+noise and noise trials (Figure 6C), and (iv) interaction between TPR and stimulus (Figure 6D), whereby the interaction map (iv) was computed as the difference between the two difference maps (ii) and (iii).

We verified that the results from Figure 8 and Figure 8–figure supplement 1 were not affected by baseline fluctuations caused by variations in ITI (the analyses in Figure 5 and Figure 7 were based on stimulus or choice-specific components of the fMRI response, which are unlikely to be affected by non-specific baseline response fluctuations). In a control analysis, we repeated the brainstem response estimates from Figure 8 after excluding 50% of all trials with the shortest ITIs. We also verified that the results were not due to non-linearity (i.e., floor and ceiling effects) in the TPR measurements. To this end, we repeated the analyses from Figure 8 after removing the 20% of trials with extreme overall pupil size during the trial (top and bottom 10% pupil measurements in the TPR-window, without subtracting the pre-trial baseline pupil). Both sets of control analyses yielded qualitatively identical results as in Figure 8, and Figure 8–figure supplement 1 (data not shown).

We computed the correlation between TPR and task-evoked fMRI responses across all trials (Figure 8F) after first removing (via linear regression) effects of signal presence from both TPR and the task-evoked fMRI responses. These correlations were computed separately for each individual subject and then tested against 0 across the group (see section *Statistical comparisons*).

To compute the "combined neuromodulatory brainstem signal" (Figure 9, Figure 9–figure supplement 1), we performed a multiple linear regression against TPR for the following ROIs: LC region, VTA, SN, BF-sept and BF-subl (see section *Definition of regions of interest*). This yielded a subject-specific set of weights per ROI that maximized the correlation to TPR. We then used these weights to compute, for each subject, a linear combination of the single-trial ROI-responses. The resulting signal was z-scored, with the purpose of using it as a predictor variable within the DDM framework (see section *Analysis and modeling of choice behavior*). Likewise, the "combined colliculi signal" was computed based on a linear combination of responses in SC and IC.

Quantification of orientation-specific responses in early visual cortex Orientation-specific response to the low-contrast target gratings in visual cortical areas V1-3 (Figure 5B; see section Definition of regions of interest) were computed using multi-voxel pattern analysis with a leave one out cross-validation procedure, as follows. First, per scanning session, we selected the 100 voxels with the strongest orientation preference from the unconstricted "center" sub-region that corresponded retinotopically to the target stimulus. To this end, we selected the 50 most positive and 50 most negative t-values for the comparison between clockwise (CW) vs. counter-clockwise (CCW) stimuli on signal+noise trials:

$$t = \frac{\bar{x}_{cw}}{sem_{cw}} - \frac{\bar{x}_{ccw}}{sem_{ccw}} \tag{3}$$

where subscripts *cw* and *ccw* were clockwise and counter-clockwise target signals, respectively $\bar{x}_{cw}$ and $\bar{x}_{ccw}$ were the two sample means, and sem_{cw} and sem_{ccw} were the two corresponding standard errors. Second, the responses of the selected 100 voxels were arranged as vectors, each of which constituted a multi-voxel response pattern for one trial. Third, these single-trial response patterns were separately averaged across counter-clockwise and across clockwise trials, leaving out one trial per iteration. Fourth, a "template pattern" was constructed as the difference between the mean counter-clockwise and clockwise response patterns. Fifth, the pattern of fMRI responses of the remaining trial (including noise trials) was correlated with the template pattern (on CW blocks), or with the inverted template (on CCW blocks). This procedure yielded a correlation coefficient per trial, which quantified the strength of the orientation-specific component of the population response of V1-V3.

Quantification of choice-specific responses: univariate approach We used two approaches to identify choice-encoding regions of interest (ROIs) and to quantify choice-specific signals therein. First, univariate logistic regression of binary choice against task-evoked fMRI responses (Figure 7A, and Figure 7–figure supplement 1A,B) described in this section; second, "searchlight decoding" of choice based on multivariate local fMRI response patterns (Figure 7B, and Figure 7–figure supplement 1C,D), which is described in the subsequent section.

In the first approach, the following logistic regression model was fitted to the single-trial evoked fMRI responses of each voxel:

$$P(yes) = \frac{exp(\beta_0 + \beta_1 A)}{1 + exp(\beta_0 + \beta_1 A)} \tag{4}$$

where $P(yes)$ was the predicted probability that the subject made a "yes"-choice for a given value of A (in the model, a binary vector describing the choice identity (1,"yes"; 0, "no") served as the dependent variable), A was a vector of fMRI response amplitudes, and β_0 and β_1 were the parameters of the fit. The parameters β_0 and β_1 were adjusted by an optimization routine to fit the measured probability of subjects' choices across all trials. The voxel was assigned the slope parameter, b1, which quantified the link between the task-evoked fMRI responses of that voxel and the subject's behavioral choices. Performing this analysis for all voxels yielded a map of choice-specific activity for each subject. We performed the analysis separately for signal+noise trials and for noise trials and averaged the resulting maps. That way, the effect of choice was isolated, discarding potential differences between signal+noise and noise trials.

In the main task, "yes"- and "no"-choices were mapped onto motor responses with different hands. We therefore computed the logistic regression not only for each voxel's fMRI response, but also for the lateralization—that is, the difference between a voxel's response and its homotopic counterpart in the contralateral hemisphere. The mapping between the "yes"- vs."no"-choice and response hands was counterbalanced across subjects. Thus, the lateralization reflected the effector-specific activity (left vs. right hand button presses). Because in any given session "yes"- and "no"-choices were mapped onto button presses with different hands, the lateralization was a proxy of the choice-specific activity, encoded in the format of a plan to act (e.g., see [2]). Before computing the lateralization, each voxel was flipped with its homotopic counterpart in the contralateral hemisphere for half of the subjects, for whom the mapping was: "yes" –> right hand and "no" –> left hand. The lateralization was then expressed with respect to the hand for "yes" for all subjects [26]. Consequently, lateralization values could be pooled across subjects without cancellation. The result of this analysis was visualized on the right hemisphere (Figure 7A, and Figure 7–figure supplement 1A,B).

We computed choice-specific fMRI response lateralization for a number of ROIs that were symmetric between the two hemispheres (see section *Definition of regions of interest*; Figure 7A, and Figure 7–figure supplement 1A,B). First, per scanning session, we selected the 50 voxels from the hemisphere contralateral to the hand used for reporting "yes" with the most positive t-values for the comparison between "yes"- vs. "no"-choices (same as Equation 3, but now taking the difference of "yes"- and "no"-trials), and vice versa for the ipsilateral hemisphere. This adds up to a total of 100 voxels. Second, we separately averaged the single-trial responses across each of these two groups of voxels. Third, we computed the single-trial lateralization as the difference between the mean fMRI response of the "yes" and the "no" voxel group. Performing this analysis for each trial enabled computing single-trial correlations of fMRI lateralization values with TPR and behavioral choice, respectively.

To compute the "combined lateralization signal" (Figure 7D,E), we fitted multiple logistic regression of choice on choice-specific responses in aIPS1 and IPS/PostCeS, and obtained the weighted sum of these responses that maximized the correlation to choice.

Quantification of choice-specific responses: multivariate (searchlight decoding) approach We used searchlight multi-voxel pattern classification to identify additional cortical regions which might encode the choice in other formats (e.g., more fine-grained patterns [86, 87]) than the coarse hemispheric lateralization of response amplitudes. We used the LIBSVM implementation of a linear support vector machine [119], with a standard cost value of c = 1. A sphere of voxels (i.e. "searchlight") was selected around a given voxel with a radius of 10 mm (435 voxels). From these voxels, the task-evoked fMRI responses were extracted, which made up the pattern vectors that were used for multivariate pattern classification of choices. We assigned each vector a label corresponding to the choice of the subject ("yes" vs. "no"). The pattern vectors of all but one trial were then used to train a support vector machine to predict the category of the left-out pattern. After training, we validated the model by comparing the true label of the left-out pattern with the label predicted by the model. We repeated this train-test approach iteratively for each trial and calculated a mean cross-validated precision score across all trials for this searchlight. This precision score was assigned to the center voxel of the searchlight sphere. This procedure was repeated iteratively, for all voxels in the brain, yielding a map of precision scores for each subject. We performed the choice decoding analysis separately for signal+noise and noise trials and averaged the resulting maps. This combined individual map was spatially smoothed (FWHM: 4 mm) for group-level analyses.

Choice-specific responses for ROIs defined in the searchlight analysis (see section Definition of regions of interest; Figure 7B, and Figure 7–figure supplement 1C,D) were computed as follows. First, per scanning session, we selected from each ROI the 100 voxels with the strongest choice preference. To this end, we selected the 50 most positive and 50 most negative t-values for the comparison between "yes"- vs. "no"-choices (same as Equation 3, but now taking the difference of "yes"- and "no"-trials). Second, a "template pattern" was constructed by averaging the task-evoked fMRI response of each selected voxel across all "yes" trials minus the average on all "no" trials. One trial was left out (leave one out cross-validation). Third, the pattern of fMRI responses of the remaining trial was correlated with the template pattern. This procedure yielded a correlation coefficient per trial, which quantified the sign and strength of the choice-specific pattern response. Again, performing this analysis for each trial enabled computing single-trial correlations of choice-specific pattern responses with TPR and behavioral choice, respectively.

To compute the "combined searchlight signal" (Figure 7D,E), we fitted multiple logistic regression of choice on choice-specific responses in IPL, SPL, aIPS2, pIns, PreCeS/IFG and MFG, and obtained the weighted sum of these responses that maximized the correlation to choice. Although trial-to-trial variations in RT had been removed (via linear regression) from the fMRI response of each individual voxel (see above), there was a small, but statistically significant correlation between the "combined searchlight signal" and trial-by-trial RTs (group average r = 0.027, p=0.004, permutation test). However, this correlation was significantly smaller than the correlation between the "combined searchlight signal" and behavioral choice (group average difference in correlation: "Δr=0.481, p<0.001). Thus, the multivariate choice decoder primarily decoded trial-to-trial variations in choice rather than in RT.

Receiver-operating characteristic (ROC) analysis of stimulus- and choice-specific responses We used ROC-analysis (Green and Swets, 1966) to quantify the reliability of the orientation- and choice-specific cortical responses at the single-trial level (Figure 5–figure supplement 1, Figure 7, and Figure 7–figure supplement 1). The ROC index ranged between 0 and 1, and quantified the probability with which one could predict the experimental variable of interest (in our case: signal presence or behavioral choice) based on responses measured during individual trials. An index of 0.5 implied chance level prediction. To remove effects of choice when predicting the stimulus, we averaged ROC indexes for signal presence obtained separately on "yes" and "no" trials (Figure 5–figure supplement 1). Similarly, to remove effects of stimulus when predicting choice, we averaged ROC indexes for choice obtained separately on signal+noise and noise trials (Figure 7, and Figure 7–figure supplement 1).

Analysis of stimulus independent TPR effect on cortical signals We used linear regression analysis to evaluate the hypothesis that the TPR-linked modulation of choice-specific responses is intrinsic, that is, if the modulation remained when factoring out effects of the physical evidence (Figure 7). We obtained the residual choice-specific (combined) responses by removing (via linear regression) variations due to stimulus (binary vector describing the stimulus identity (1, signal+noise; 0, noise).

Across 5 bins, defined by TPR, we then fitted the following linear model to the residual choice-specific responses:

$$CTx = \beta_0 + \beta_1 TPR \tag{5}$$

where *Ctx* was a vector of mean residual (with effects of stimulus removed, see above) choice-specific responses per bin, TPR was a vector of mean task-evoked pupillary responses per bin, and β_0 and β_1 were the free parameters of the fit.

1.4.9 Statistical comparisons

We used nonparametric paired permutation tests to test for significant differences between behavioral estimates (Figure 2, Figure 2–figure supplement 1, Figures 3 and 4, and Figure 4–figure supplement 2), task-evoked fMRI responses (Figure 7, Figure 7–figure supplement 1, Figure 8, and Figure 8–figure supplement 1), stimulus- / choice-predictive indices (Figure 5–figure supplement 1, Figure 7, and Figure 7–figure supplement 1), and regression beta weights / coefficients (Figures 1, 2, Figure 2–figure supplement 1, Figures 3, 7, 8, Figure 8–figure supplement 1, Figure 9, and Figure 9–figure supplement 1) from different trial categories, or to test them against zero (against 0.5 in the case of stimulus- / choice-predictive indices). Statistical tests were performed at the group level, using the individual subjects' mean parameters as observations. For each comparison, we randomly permuted the labels of the observations (e.g., the regressor label of the beta estimates), and recalculated the difference between the two group means (10,000 permutations). The p-value was the fraction of permutations that exceeded the observed difference between the means.

We used nonparametric permutation tests within the FSL Randomise implementation to test cluster-corrected task-evoked fMRI responses against 0 (Figure 5), linear regression coefficients against 0 (Figure 8), logistic regression beta weights against 0.5 (Figure 7, Figure 7–figure supplement 1), and searchlight classification precision scores against 0.5 (Figure 7, Figure 7–figure supplement 1). Randomise implemented 10,000 randomly generated permutations of the data to perform a Monte Carlo-style permutation test. This procedure was robust with respect to inflated false-positive rates [120]. In the majority of cases, we used a cluster correction threshold of $p<0.01$. For the logistic regression of binary choice against lateralized task-evoked fMRI responses (Figure 7, Figure 7–figure supplement 1), and in case of the brainstem TPR correlation (Figure 8F; accounting for the comparably low sensitivity of fMRI measurements in the brainstem), we used a cluster correction threshold of $p<0.05$.

1.4.10 Definition of regions of interest (ROIs)

The analyses of fMRI signals focused on a number of ROIs, which were defined in each individual brain using a variety of criteria described in this section.

Brainstem nuclei We defined the following brainstem nuclei (Figure 8, and Figure 8–figure supplement 1) based on anatomical criteria.

Locus coeruleus (LC) The LC was delineated, separately for each subject and scanning session, by means of a specific (TSE) high-resolution structural MRI scan (Figure 8A, and Figure 8–figure supplement 1A). The LC, as some other brainstem structures, has increased neuromelanin concentration and can thus be identified in the TSE scans as two bilateral local maxima of elevated signal intensity within the midbrain tegmentum and pons, at the floor of the fourth ventricle [97, 98, 107]. We delineated LC voxels in the slice with the clearest signal intensity elevation in the designated region and in the one slice above and one below. The ROI definition was initially performed by the first author and then independently verified by another author. For the analysis of functional LC responses, we transformed the individually delineated high-resolution LC ROI to subject- and session-specific EPI space (trilinear affine transformation, using FSL FLIRT) after first (i) co-registering the high-resolution partial field-of-view TSE scan and whole-brain TSE scan (with 6 degrees of freedom), (ii) co-registering the whole-brain TSE scan and the T2-weighted anatomical scan (12 degrees of freedom), and (iii) concatenating the estimated transformation matrices. We then determined the minimum number of EPI voxels (obtained at lower resolution) for which the probability of the anatomical LC ROI contained in this EPI voxel was larger than 0 within each of the 14 subjects. The average number of voxels with a non-zero probability of containing the anatomical LC ROI was 18 voxels, and the smallest number across the group was 12. We therefore included for each subject the 12 voxels with the largest probability of containing the LC. The LC region time series was computed as the average of the time series from these 12 voxels, weighted by their probability values. For comparison, we used the same procedure, including the same number of voxels, for the other brain-

stem regions. For the most spatially specific definition of the LC region afforded by our measurements we computed a weighted average across the two voxels with the largest probability of containing the LC (Figure 8D).

Superior and inferior colliculi (SC and IC) We manually delineated the SC and IC based on the anatomical standard (MNI) brain as the two elevated "hills" at the dorsal part of the midbrain.

Other brainstem nuclei We used probabilistic atlases in anatomical standard (MNI) space to delineate the following other brainstem nuclei: ventral tegmental area (VTA), substantia nigra (SN) [121], and two divisions of the basal forebrain (BF): one that includes cell groups within the septum and the horizontal limb of the diagonal band (BF-sept), and one that includes the sublenticular part of the basal forebrain (BF-subl). The probabilistic BF atlases [89] were obtained from the SPM anatomy toolbox [122]. As we did for the LC, we transformed these anatomical ROIs to subject- and session-specific EPI space and thresholded the resulting masks such as to include only the 12 voxels with the largest probability of containing the particular structure. Likewise, ROI-level time series were computed as the average of the time series from these 12 voxels, with the weights corresponding to each voxel's probability value.

Anterior cingulate cortex We parcellated the cortex with the Freesurfer Destrieux atlas [123], and obtained anterior cingulate cortex (ACC) masks by transforming the label G_and_S_cingul-Mid_Ant_d to subject- and session-specific functional space (with Freesurfer mri_label2vol tool).

Visual cortical areas V1-V3 Stimulus-responsive and surrounding sub-regions of early visual cortical areas V1-V3 (Figure 5) were defined in two steps. First, the boundaries between these visual cortical areas were identified by retinotopic mapping via population receptive field imaging (i.e., quantifying, for each voxel, the preferred polar angle and eccentricity) [124]. To this end, in a separate scanning session, we presented moving bar apertures consisting of a large number of randomly oriented Gabor patches (100% contrast). The width of the bar subtended 25% of the stimulus radius. Four bar orientations (0°, 45°, 90°, 135°) and two different motion directions for each bar were used, yielding a total of eight different bar configurations within a given scan. Subjects fixated at the center of the screen, and pressed a button with their right index finger when the fixation dot changed color (from red to green) at random time intervals. Second, using a localizer run in each session of the main experiment, we identified the sub-regions of V1-V3 that corresponded retinotopically to the stimulus judged in the main experimental (yes-no) task. To this end, full-contrast gratings were presented at the same position as the stimuli in the main experimental task, in periodic alternation with central fixation of an otherwise blank screen (block duration: 12 s). Figure 5A shows the map of localizer responses in one example subject.

Choice-specific cortical regions Choice-specific ROIs (Figure 7, and Figure 7–figure supplement 1) were defined as follows: univariate logistic regression on lateralized task-evoked fMRI responses

and multivariate searchlight decoding (see section Analysis of MRI data). Group-level cluster correction yielded the conjunction maps shown Figure 7A,B, and Figure 7–figure supplement 1A–D. Performing these analyses separately for each of the two scanning sessions allowed us to identify robust choice-specific ROIs as clusters of voxels, which exhibited statistically significant choice-specific activity within both independent sessions (conjunction maps in Figure 7A,B).

The first approach yielded the following choice-specific ROIs: hand region of primary motor cortex in the central sulcus (M1), posterior parietal cortex, in the junction of the intraparietal sulcus and the postcentral sulcus (IPS/PostCeS), and the anterior intraparietal sulcus (aIPS1) (Figure 7A, and Figure 7–figure supplement 1A,B).

The second approach yielded the following choice-specific ROIs: the superior parietal lobule (SPL), the inferior parietal lobule (IPL), a second region within the anterior intraparietal sulcus (aIPS2), the posterior insula (pIns), the junction of the precentral sulcus and the inferior frontal gyrus (Pre-CeS/IFG), and the medial frontal gyrus (MFG) (Figure 7B, and Figure 7–figure supplement 1C,D).

1.4.11 Data and code sharing

The data are publicly available on Figshare (https://doi.org/10.6084/m9.figshare.4806562; https://doi.org/10.6084/m9.figshare.4806559). Analysis scripts are publicly available on Github (https://github.com/jwdegee/2017_eLife; https://github.com/jwdegee/2014_PNAS).

1.5 Acknowledgements

We thank Daniel Lindh for help with the data collection of the main experiment. We thank Matthew McGinley for providing the stimuli, and Danielle Rijkmans, Guusje Boomgaard and Christopher David Riddell for help with the data collection for the auditory detection task. We thank Konstantinos Tsetsos, Angela Yu, David J Heeger, Birte U Forstmann and Todd Hare for discussion. This research was supported by the German Research Foundation (DFG, SFB 936/Z1 and DO 1240/3–1, to THD), the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 604102 (Human Brain Project, to THD), and the European Research Council (ERC) under grant agreement no. 283314-NOREPI (to SN).

1.6 Supplementary Figures

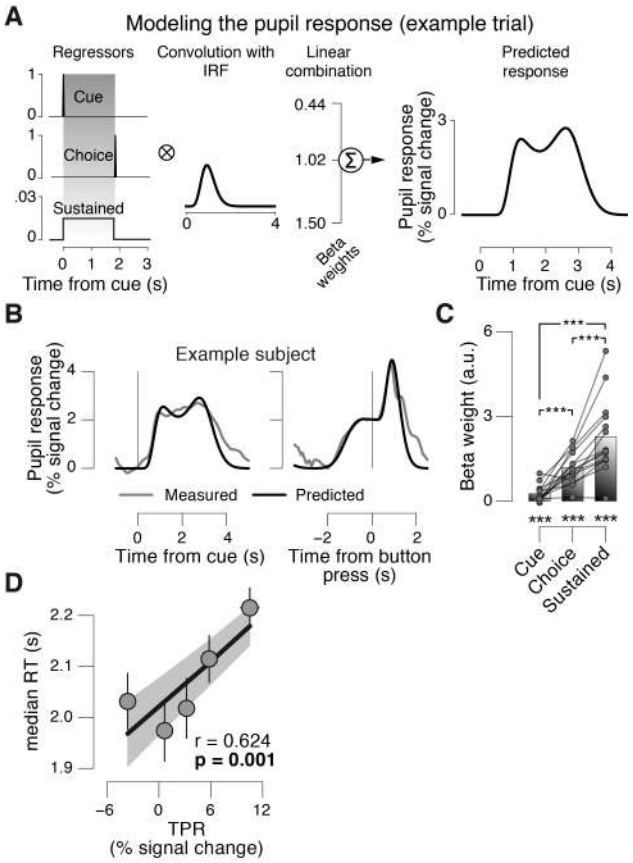


Figure S1: [Figure 1–figure supplement 1] Linear modeling of TPR. The peripheral system (i.e., nerves and the muscles of the iris [110]) that transforms central neural inputs originating from the brainstem into TPR is sluggish and acts as a low-pass filter [41, 109]. **(A)** Linear modeling of TPR (see Materials and methods). We used a previously established general linear model (GLM) to estimate the relative contribution of three putative underlying neural input components (see Materials and methods; [40]). Gray-shaded interval, decision interval (between cue, i.e., decision onset, and button press). The three beta weights are the best fitting parameter estimates for the subject from panel B. IRF, impulse response function. **(B)** TPR time course from example subject, aligned to cue (left) and button press (right). Gray line, mean TPR (n = 624 trials). Black line, model prediction. The model provided a good fit of the shape of the TPR time course in most subjects. **(C)** Mean beta weights for all temporal components. As observed previously [40], the predominant input (i.e., largest beta-weight) was a sustained component that spanned the interval between cue- and choice-components. Data points, individual subjects; ***p<0.001. **(D)** Correlation between TPR and reaction time (RT) (5 bins). Shading or error bars, s.e.m. All panels: Group average (N = 14); stats, permutation test.

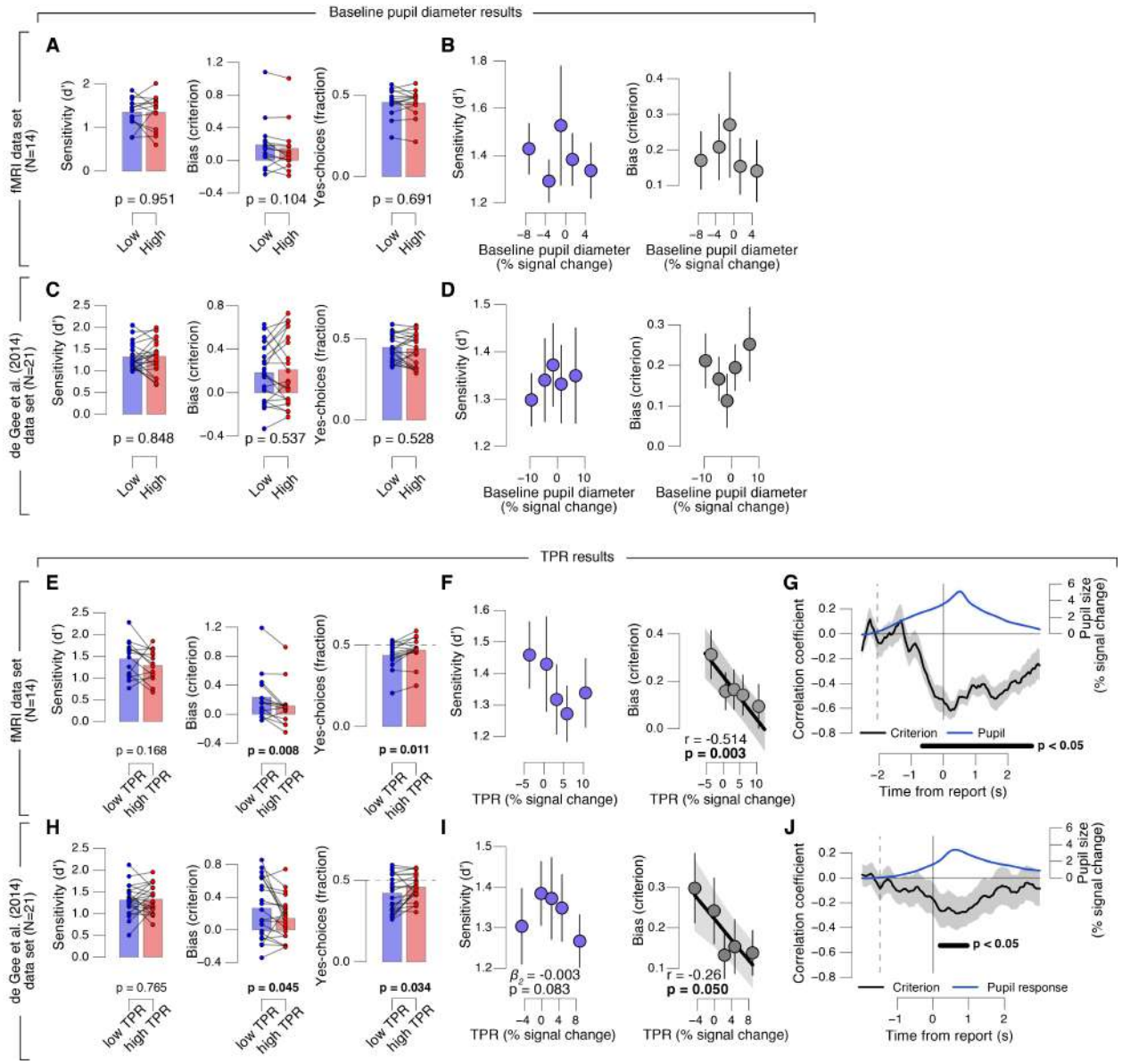


Figure S2: [Figure 2–figure supplement 1] Phasic arousal predicts reduction of choice bias. (A–D) As main Figure 2 panels A, B, D, E, but now after splitting trials by (or as a function of) baseline pupil diameter, quantified as the mean pupil size in the 0.5 s before decision interval onset. **(E–J)** As main Figure 2 panels A-F, but now without removing trial-to-trial variations from TPR due to reaction time (RT). The pattern of results is qualitatively identical to the one in Figure 2. All panels: group average (N = 14); shading or error bars, s.e.m.; stats, permutation test. Panels A, C, E, H: data points, individual subjects.

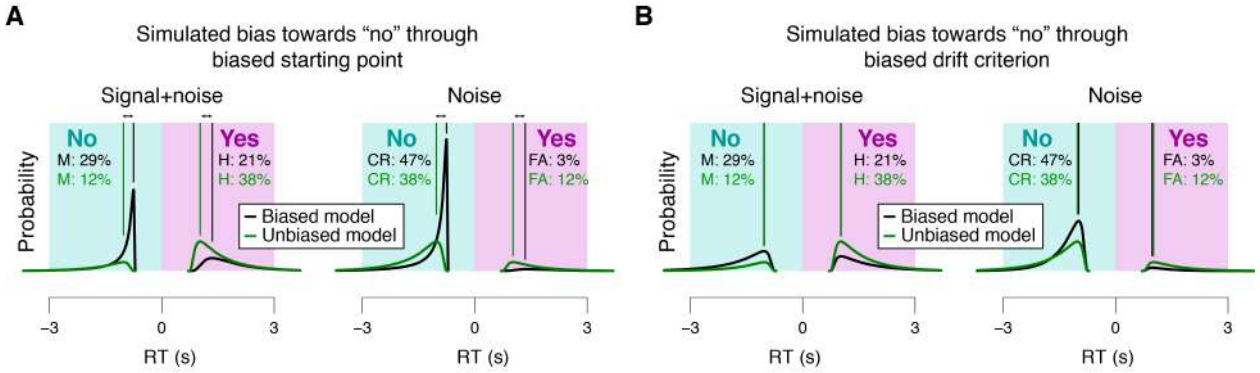


Figure S3: [Figure 4–figure supplement 1] Effects of starting point vs. drift criterion on RT distributions. Analytical RT distributions sorted by the four SDT categories: "yes"- and "no"-choices, as well as signal+noise and noise trials. Within each category, RT distributions are shown separately for a biased model (black) and an unbiased model (green), whereby biased model refers to a model producing unequal fractions of "yes"- and "no"-choices. RTs for "no"-choices were sign-flipped for illustration purposes. A conservative choice bias was implemented in two separate mechanisms in the two panels, both producing the same change on the fraction of choices. These mechanisms have distinguishable effect on the shape of the RT distributions, most evident in the mode. **(A)** Conservative choice bias through setting the starting point of the accumulation process closer to the "no"-bound. Straight lines, mode (i.e., maximum) of the analytical RT distributions. A shift in starting point produced a substantial shift of the mode of the RT distribution (compare black and green vertical lines). **(B)** Conservative choice bias through changing drift criterion in the direction of the "no"-bound. This shift in drift criterion had a negligible effect on the mode of RT distributions (green and black vertical lines are on top of one another).

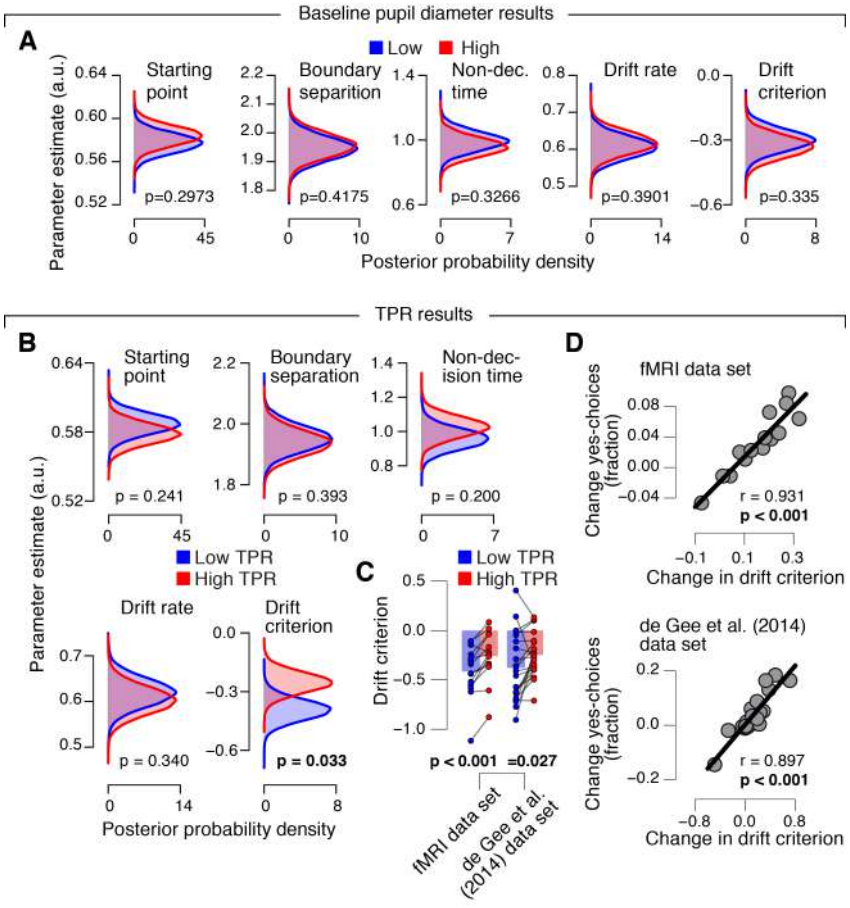


Figure S4: [Figure 4–figure supplement 2] Phasic arousal predicts reduction of accumulation bias. (A) As main Figure 4C, but now after splitting trials by baseline pupil diameter, quantified as the mean pupil size in the 0.5 s before decision interval onset. **(B–D)** As main Figure 4C–E, but now without removing trial-to-trial variations from TPR due to reaction time (RT). The pattern of results is qualitatively identical. Data points, individual subjects; stats, permutation test.

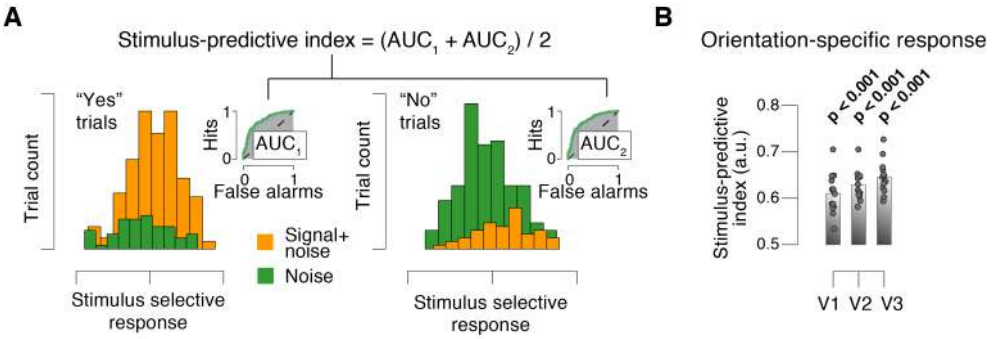


Figure S5: [Figure 5–figure supplement 1] Quantifying single-trial reliability of stimulus-specific responses. (A) The stimulus-predictive index was calculated from the receiver operating characteristic (ROC) curve, separately for "yes"- (left) and "no"-choices (right). Each panel shows distributions (taken from area V1 of one example subject) of single-trial pattern responses (see Materials and methods) for signal+noise and noise conditions. ROC curves (insets) were constructed by shifting a criterion across both distributions, and plotting against one another, for each position of the criterion, the fraction of trials for which responses were larger than the criterion. The area under the resulting ROC curve (AUC; grey shading), here referred to as "predictive index", quantifies the probability with which an ideal observer can predict the signal presence from the single-trial fMRI response. Predictive indices calculated separately for "yes"- and "no"-trials were pooled (averaged) into a single stimulus-predictive index. The predictive index in this example was 0.63. **(B)** Stimulus-predictive indexes for orientation-specific responses (signal+noise vs. noise, irrespective of choice; see Materials and methods and panel A). Data points, individual subjects; stats, permutation test.

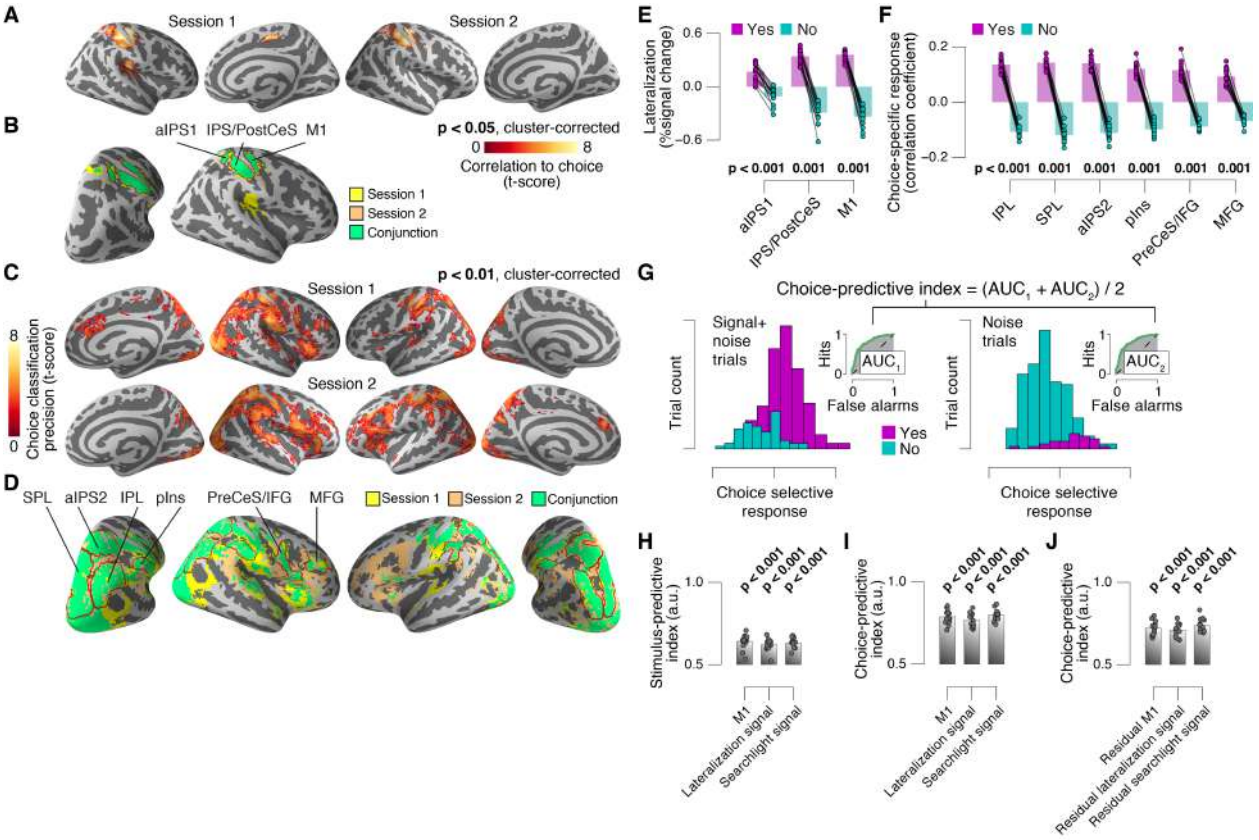


Figure S6: [Figure 7–figure supplement 1] Identifying choice-specific cortical signals. (A) Session-wise maps map of logistic regression coefficients (choice against fMRI lateralization). **(B)** Overlay of both maps from panel A, to identify robust and replicable choice-specific responses. **(C)** Session-wise maps of searchlight choice classification precision scores. **(D)** Overlay of both maps from panel C, to identify robust and replicable choice-specific responses. A-D: all tested against 0.5 at group level; red outlines, ROIs used for further analyses. **(E,F)** Choice-specific responses in ROIs from panels B and D, respectively, separately for "yes"- and "no"-choices. **(G)** Quantifying the reliability of choice-specific cortical responses, using single-trial, lateralized M1-responses of an example subject. As Figure 5–figure supplement 1A, but now for prediction of choice, rather than of signal presence. To remove effects of signal presence, we averaged ROC indexes for choice computed separately for signal+noise and noise conditions. The resulting measure is analogous to "choice probability" employed in monkey electrophysiology. The predictive index in this example was 0.82. All panels: group average (N = 14); data points, individual subjects; stats, permutation test. aIPS, anterior intraparietal sulcus; IPS, intraparietal sulcus; PostCeS, postcentral sulcus; M1, primary sensorimotor cortex (hand area); SPL, superior parietal lobule; IPL, inferior parietal lobule; pINS, posterior insular cortex; PreCeS, precentral sulcus; IFG, inferior frontal gyrus; MFG, medial frontal gyrus. **(H)** Stimulus-predictive indexes for choice-specific responses ("yes" vs. "no", without first taking out effects of stimulus as in panel G). "Combined lateralization" signal, weighted sum of choice-specific responses across ROIs obtained through mapping significant lateralization with respect to the hand movement (except M1; from Figure 7A; see Materials and methods). "Combined searchlight" signal, weighted sum of choice-specific responses across ROIs obtained through searchlight classification (from Figure 7B). **(I)** As panel H, but for choice-predictive indexes (signal+noise vs. noise, without first taking out effects of choice as in Figure 5–figure supplement 1A). **(J)** As panel I, but after removing effects of physical stimulus via linear regression (see Materials and methods). All panels: group average (N = 14); data points, individual subjects; ***p<0.001; stats, permutation test.

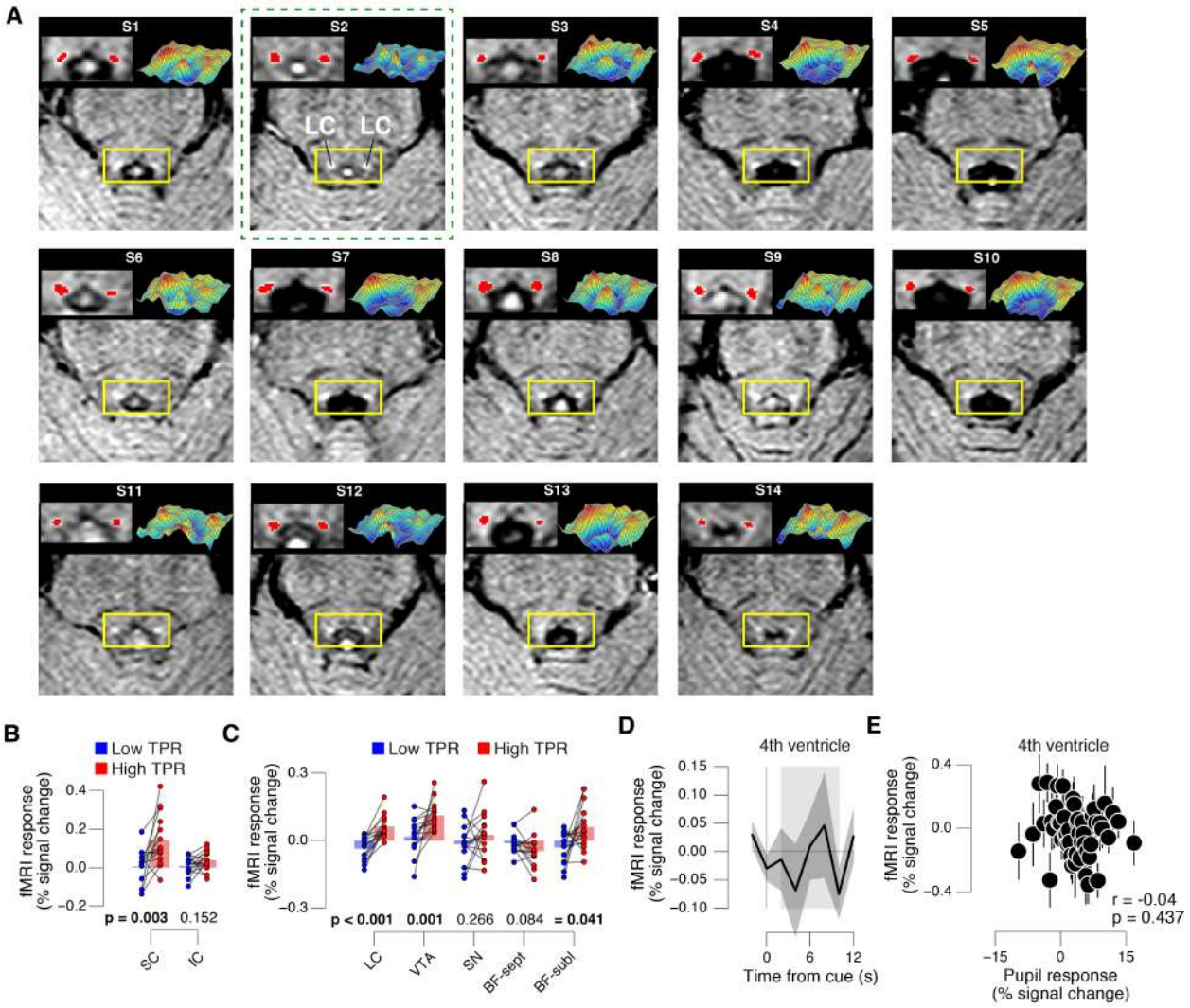


Figure S7: [Figure 8–figure supplement 1] TPR-linked brainstem responses. (A) As Figure 8A for all subjects. Green outline, subject presented in Figure 8A. **(B)** Task-evoked responses in SC and IC, separately for low and high TPR trials. **(C)** As panel B, but for LC, SN, VTA, BF-sept, and BF-subl. Data correspond to Figure 8 panels B and E, but is now represented as bar graphs of task-evoked fMRI response scalar measures (see Materials and methods). **(D)** Remaining task-evoked fMRI response in the fourth ventricle (after RETROICOR). The fourth ventricle ROI was delineated in the TSE scan, and its response computed by averaging across all voxels covering the ventricle. Gray-shaded interval, task-evoked fMRI response measure time window. Residual signal fluctuations in the fourth ventricle were uncorrelated to task events. Shading, s.e.m. **(E)** Task-evoked fMRI responses in the fourth ventricle plotted against TPR. Data were binned for display (50 bins). Panels B–E: group average (N = 14). Panels B,C: data points, individual subjects; stats, permutation test.

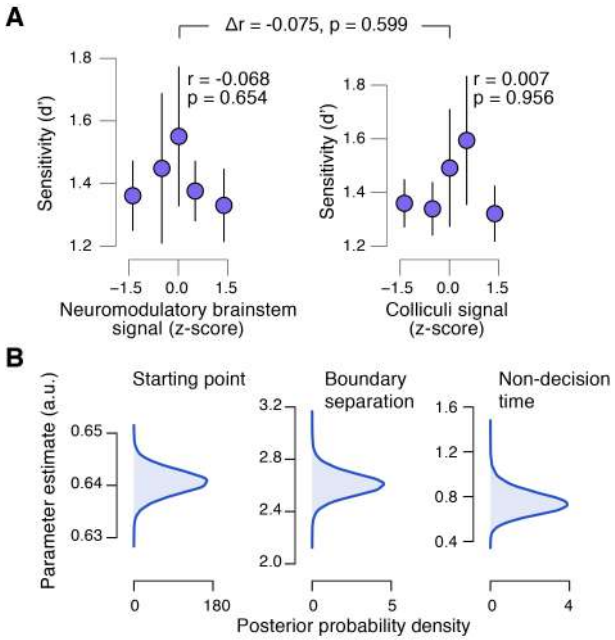


Figure S8: [Figure 9–figure supplement 1] Brainstem responses are not associated to sensitivity. (A) Correlation between perceptual sensitivity (d') and the combined neuromodulatory brainstem response (linear combination of responses in LC, SN, VTA, BF-sept, and BF-subl maximizing the correlation to TPR; see Materials and methods), and the combined colliculi response (linear combination of responses in SC and IC maximizing the correlation to TPR) (5 bins). Stats, permutation test. **(B)** Group-level posterior probability densities for means of the remaining parameters in the DDM regression model, through which we assessed the linear relationship between the trial-by-trial drift and the trial-by-trial combined neuromodulatory response and the combined colliculi response (see Materials and methods). All panels: group average (N = 14); shading or error bars, s.e.m.

Status: in review
de Gee JW, Tsetsos K, Schwabe L, Urai AE, McCormick DA, McGinley MJ* & Donner TH* (2019). *bioRxiv.*
(*=Shared senior authorship)

2

Phasic arousal suppresses suboptimal decision biases in mice and humans

Conceptualization by JWG, KT, LS, DAM, MJM, THD; Investigation by JWG, KT, MJM; Formal analysis by JWG, KT, AEU, MJM; Writing–original draft by JWG, KT, MJM, THD; Writing–review and editing by JWG, KT, LS, AEU, DAM, MJM, THD.

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.

2.1 Introduction

The global arousal state of the brain changes from moment to moment [6, 11]. 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 [6–10]. These same arousal systems are phasically recruited during elementary decisions, in relation to key computational variables such as uncertainty and surprise [6, 29–31, 45, 74, 125, 125–127].

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 [128]. Influential theoretical accounts propose that phasic arousal has an adaptive function that serves to optimize inference and choice behavior [6, 30]. 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 [20]. 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 [18, 20, 22, 23]. 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 [6]. We asked whether and how phasic arousal might also shape the preceding evidence accumulation.

Second, while brainstem arousal systems are homologously organized across mammals [129, 130], 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 [68]. 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 [20], 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 [131], or from abstract (e.g. numerical) quantities [132]. 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 [24]. 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 [76, 78, 79]. 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 [20] is shaped by task-evoked neuromodulatory responses.

2.2 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.

2.2.1 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).

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) [108, 133] (Materials and Methods), and (iii) to most specifically track noradrenergic activity [39]. 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 [134, 135].

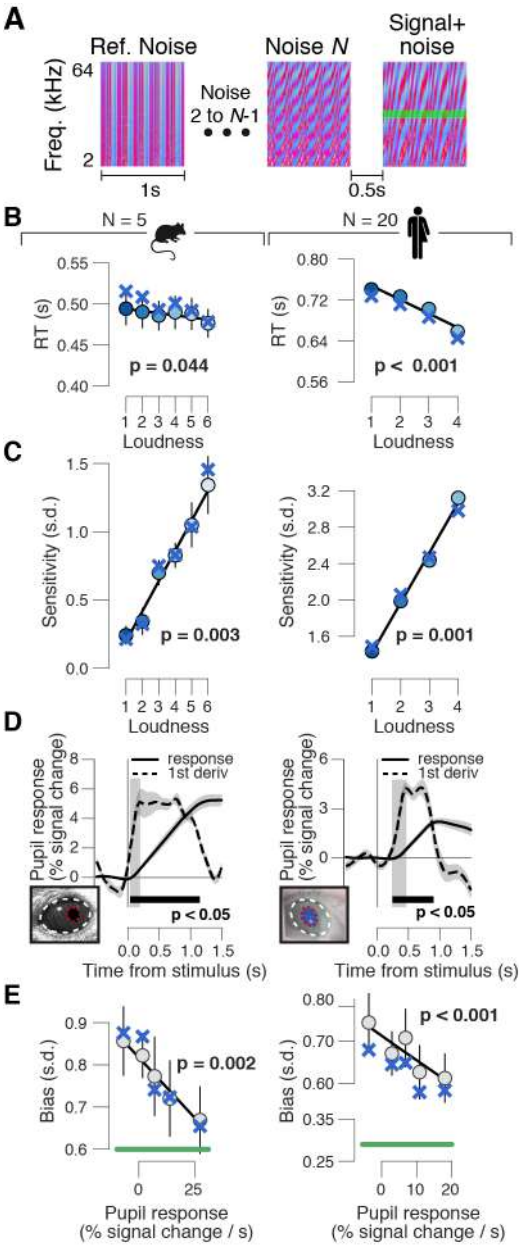


Figure 1: High phasic arousal is associated with reduced perceptual choice bias. **(A)** Auditory go/no-go tone-in-noise detection task. Schematic sequence of stimuli during a trial. Subjects were trained to respond to a weak signal (pure tone) in noise and withheld a response for noise-only sound tokens. 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.

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 [6, 136], even in the same mice dataset analyzed here [76]. 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).

2.2.2 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 [2, 18, 19, 25] 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 [25, 137]. 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

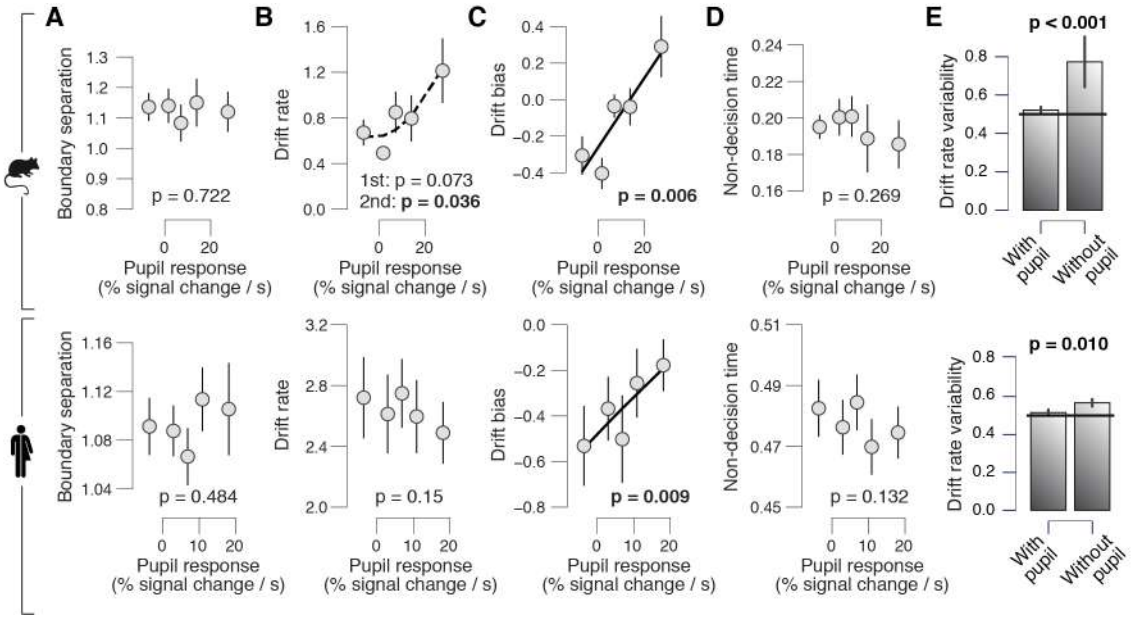


Figure 2: Phasic arousal accounts for substantial variability in sensory evidence accumulation. (A) Relationship between boundary separation estimates and task-evoked pupil response in mice (top) and humans (bottom), collapsed across loudness. See Fig. S2 for parameter estimates separately per loudness. 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.

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 [138]. 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 [18, 25]. 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 [45, 74, 125]. 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.

2.2.3 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 [33, 40, 42, 133]. The sustained component might entail motor preparatory activity [26]. 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,

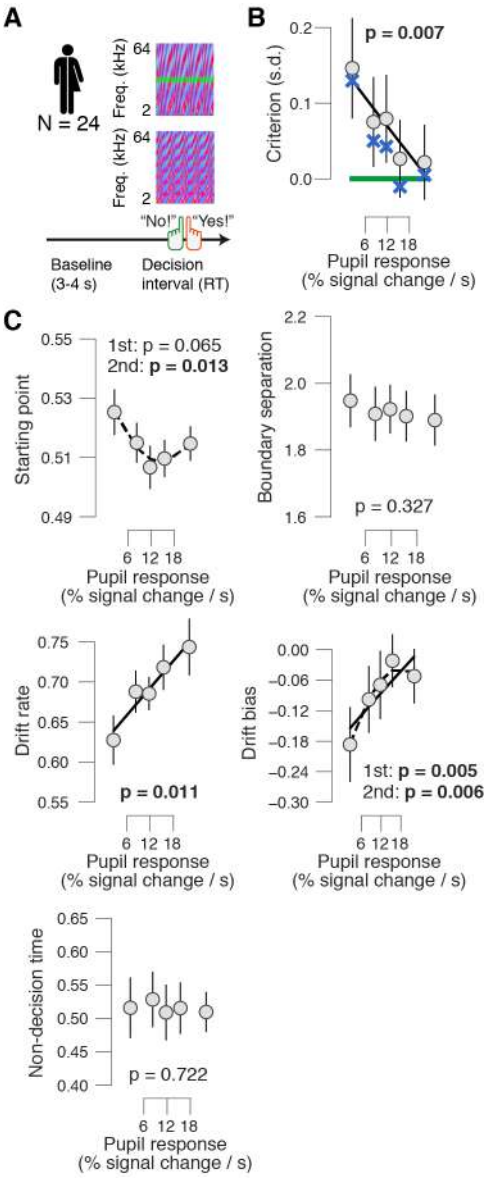


Figure 3: High phasic arousal associates with reduced perceptual evidence accumulation bias during yes/no decisions. (A) Auditory yes/no (forced choice) tone-in-noise detection task. Schematic sequence of events during a trial. Subjects reported the presence or absence of a faint signal (pure sine wave) 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.

only the changes of drift bias explained the performance-optimizing reductions in decision bias.

2.2.4 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 [24]. 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.

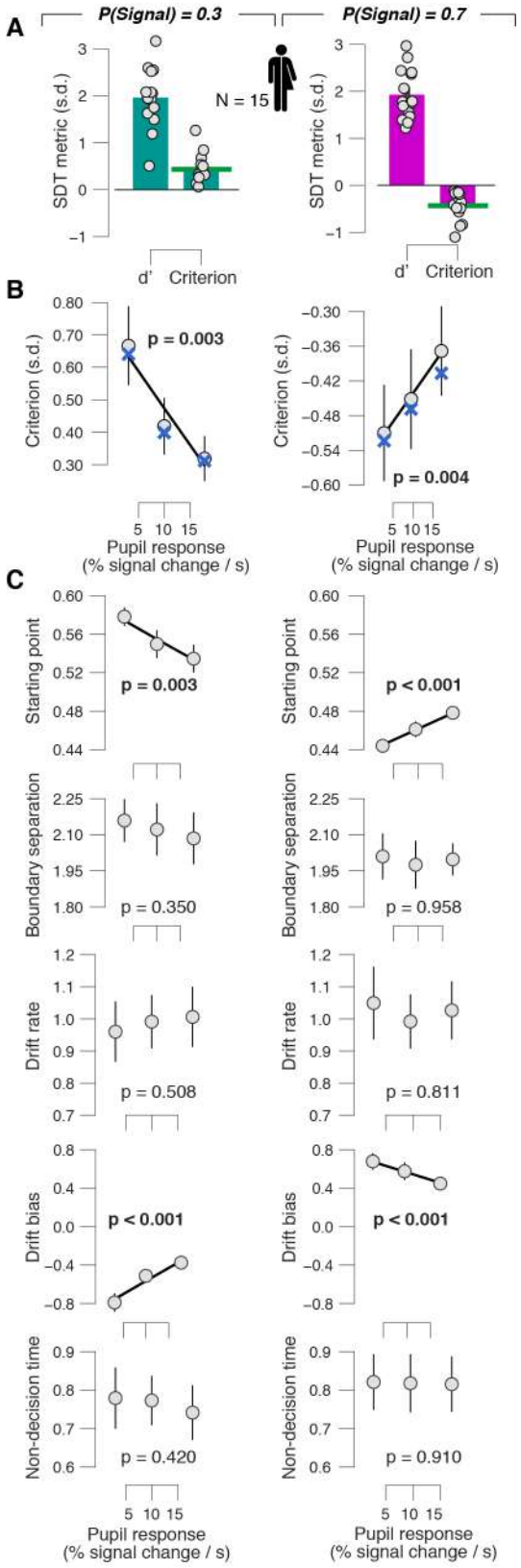


Figure 4: Phasic arousal flexibly reduces both conservative and liberal accumulation biases. (A) Overall sensitivity and choice bias (quantified by signal detection d' and criterion, respectively) in the rare signal condition (P(Signal) = 0.3; left) and frequent condition (P(Signal) = 0.7; right). Green 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.

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.

2.2.5 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 [131, 139]. 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 [140], 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: Δr = -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.

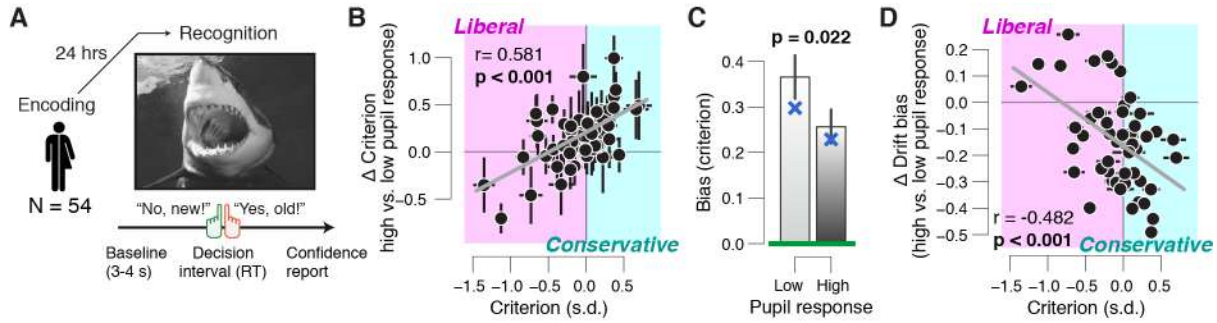


Figure 5: Phasic arousal predicts a reduction of memory accumulation biases. **(A)** A Picture yes/no (forced-choice) recognition task. Schematic sequence of events during a trial. Subjects judged whether they had seen pictures twenty-four hours previously during an encoding task (Materials and Methods). 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.

2.2.6 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 [132, 141, 142]. 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 [132, 141]. 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 sam-

ples (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 [132, 141], 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 [132]. 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; [18]); the leaky competing accumulator model (LCAM; [53]); and the leaky selective accumulator model (LSAM; [141]). 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 [141].

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 [143]. Lower BIC values indicate a model that better explained the data, whereby BIC differences of 10 indicate a decisively better fit [143]. 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

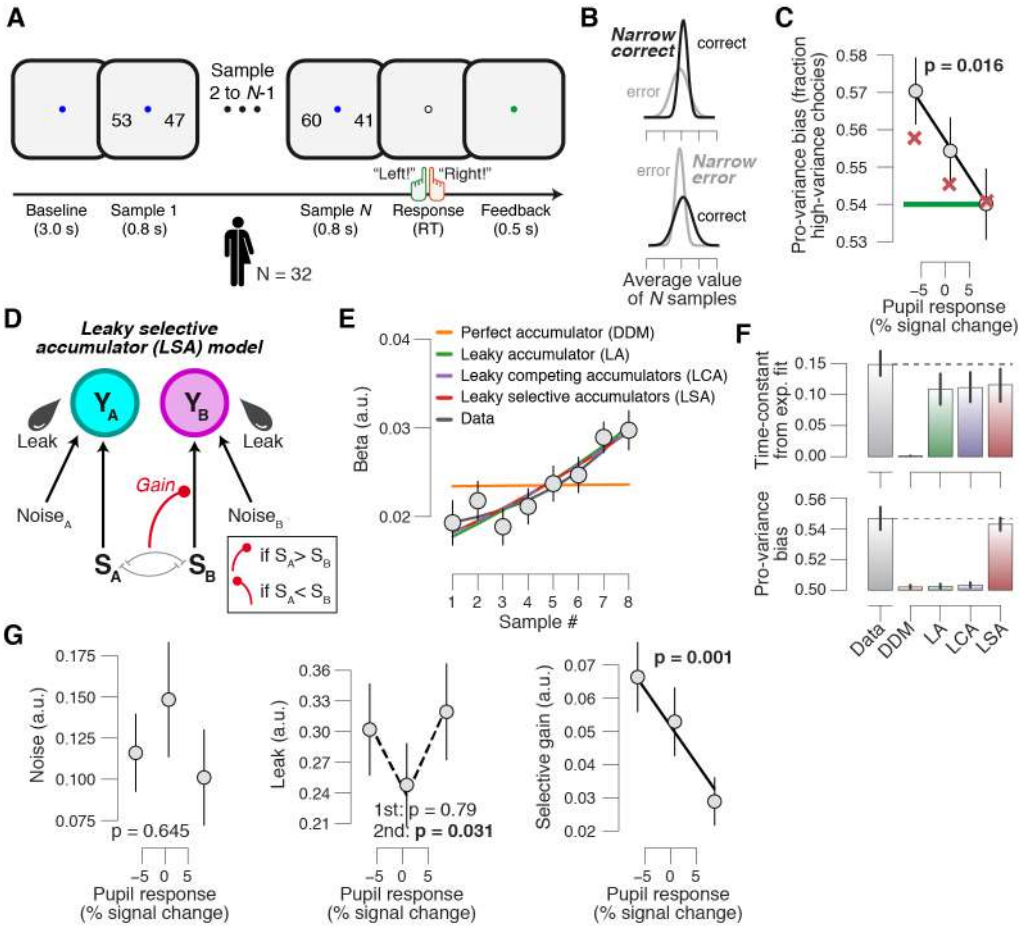


Figure 6: Phasic arousal predicts reduced risk seeking bias in a value-based choice task. (A) Value-based choice task. Schematic sequence of events during a trial. Subjects integrated two value sequences (5–8 pairs) and judged which sequence had on average the higher value. **(B)** On "narrow-correct" 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.

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 [141]. 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.

2.2.7 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 [40, 72], 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 [76], 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.

2.3 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 [6, 11]). 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 [33, 40, 74, 125, 126]. 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 [20] 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 [11, 32]. Indeed, recent work has shown that pupil diameter closely tracks several established measures of cortical arousal state [11, 32]. Changes in pupil diameter are also associated with locus coeruleus (LC) responses in humans [33, 34], monkeys [35, 36], and mice [37–39]. 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 [33, 35, 39]. 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 [25, 137], and (iii) is, under certain parameter regimens, equivalent to a reduction of biophysically detailed neural circuit models of decision-making [18, 144]. 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 [26, 52]. Third, in the go/no-go task, that subjects set an implicit boundary for no-choices [137]. 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; [76, 145]. 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) [76]. 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 [39].

Recent findings indicate that intrinsic behavioral variability is increased during sustained ("tonic") elevation of NA levels, in line with the "adaptive gain theory" [6]. 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 [146]. Second, chemogenetic stimulation of the LC in rats performing a patch leaving task increased decision noise and subsequent exploration [147]. Third, pharmacologically reducing central noradrenaline levels in monkeys performing an operant effort exertion task parametrically increased choice consistency [148]. 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 [10]. 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 [76], 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 [6]. 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 [57, 59–61, 149]. 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 [62, 67]. Neuromodulators might reduce the weight of this prior in the inference process [57, 105], thereby reducing choice biases. Another possibility is neuromodulator release might scale with uncertainty about the incoming sensory data [57, 105]. 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 [6, 91]. 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 [30]. 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 [31] and override the default state [30], 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 [150]. For example, acute stress reduces risk-seeking when making decisions involving financial gains [151], it increases overexploitation in sequential foraging decisions [152], and it impairs "model-based" goal-directed choice behavior [153, 154]. 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 [155]. 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 en-

gaged as strongly as possible? There are several possible reasons. First, phasic LC responses depend in a non-monotonic fashion on baseline LC activity [6]. 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) [6, 11]. 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 [156–158] and conveyed to brainstem arousal systems via top-down projections [6, 37]. 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 [68]. Here, we established that the shaping of evidence accumulation by phasic arousal is also governed by a conserved principle.

2.4 Materials and Methods

2.4.1 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 [76]. 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 Cen-

ter 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 [76] and two human data sets [33, 159]. Bergt et al. (2018) [159] 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 [76]), 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.

2.4.2 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 [76], 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 [40].

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 [76], 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 [159] 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 (0.34°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 0.66°wide and 0.95°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.

2.4.3 Pupil data acquisition

The mouse pupil data acquisition is described elsewhere [76]. 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.

2.4.4 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 [33, 40]. 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 [40, 133]. 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 [39] 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. $\sim$ 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.

2.4.5 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) [24] 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:

$$S(Z) - N(Z) = 0 \tag{1}$$

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

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

$$C = (0.5 \times d') + c \tag{2}$$

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:

$$\textit{Overall bias} = C - Z \tag{3}$$

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 [77] 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 [77]. Fitting the model to RT distributions for the separate responses (termed "stimulus coding" [77]) 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 [77, 160, 161]. 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. In-

dividual 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 $\hat{R}$ statistic (which compares within-chain and between-chain variance) and checked that all group-level parameters had an $\hat{R}$ 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 [137]. 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:

$$Y_A(t) = Y_A(t-1) + S_A + \xi \cdot \zeta_A(t) \tag{4a}$$
$$Y_B(t) = Y_B(t-1) + S_B + \xi \cdot \zeta_B(t) \tag{4b}$$

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:

$$Y_A(t) = (1-\lambda) \cdot Y_A(t-1) + S_A + \xi \cdot \zeta_A(t) \tag{5a}$$
$$Y_B(t) = (1-\lambda) \cdot Y_B(t-1) + S_B + \xi \cdot \zeta_B(t) \tag{5b}$$

Relative to the diffusion model, the LAM has one extra parameter, λ which is the leak of the accumulation process. For $\lambda = 1$ 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:

$$Y_A(t) = max(0, (1-\lambda) \cdot Y_A(t-1) + S_A + \xi \cdot \zeta_A(t)) \tag{6a}$$
$$Y_B(t) = max(0, (1-\lambda) \cdot Y_B(t-1) + S_B + \xi \cdot \zeta_B(t)) \tag{6b}$$

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

$$Y_A(t) = (1-\lambda) \cdot Y_A(t-1) + I_A + \xi \cdot \zeta_A(t) \tag{7a}$$
$$Y_B(t) = (1-\lambda) \cdot Y_B(t-1) + I_B + \xi \cdot \zeta_B(t) \tag{7b}$$

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:

$$I_A(t) = \theta(S_A(t), S_B(t)) \cdot S_A \tag{7c}$$
$$I_B(t) = \theta(S_B(t), S_A(t)) \cdot S_B \tag{7d}$$

with function θ as follows:

$$\theta(x, y) = \frac{1}{1 + \exp -w(x - y} \tag{7e}$$

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 [162].

We fitted the four models above using a maximum likelihood approach together with Bayesian BIC comparisons [141]. 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 [132, 141], 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 [141]). 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).

2.4.6 Statistical comparisons

We used a mixed linear modeling approach implemented in the R-package lme4 [163] 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 [111], 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:

$$\text{Model 1: } y \sim \beta_0 1 + \beta_1 S \tag{8a}$$

$$\text{Model 2: } y \sim \beta_0 1 + \beta_1 S + \beta_2 TPR^1 \tag{8b}$$

$$\text{Model 3: } y \sim \beta_0 1 + \beta_1 S + \beta_2 TPR^1 + \beta_3 TPR^2 \tag{8c}$$

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 [164]. 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* [165].

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.

2.5 Acknowledgements

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).

2.6 Supplementary Figures

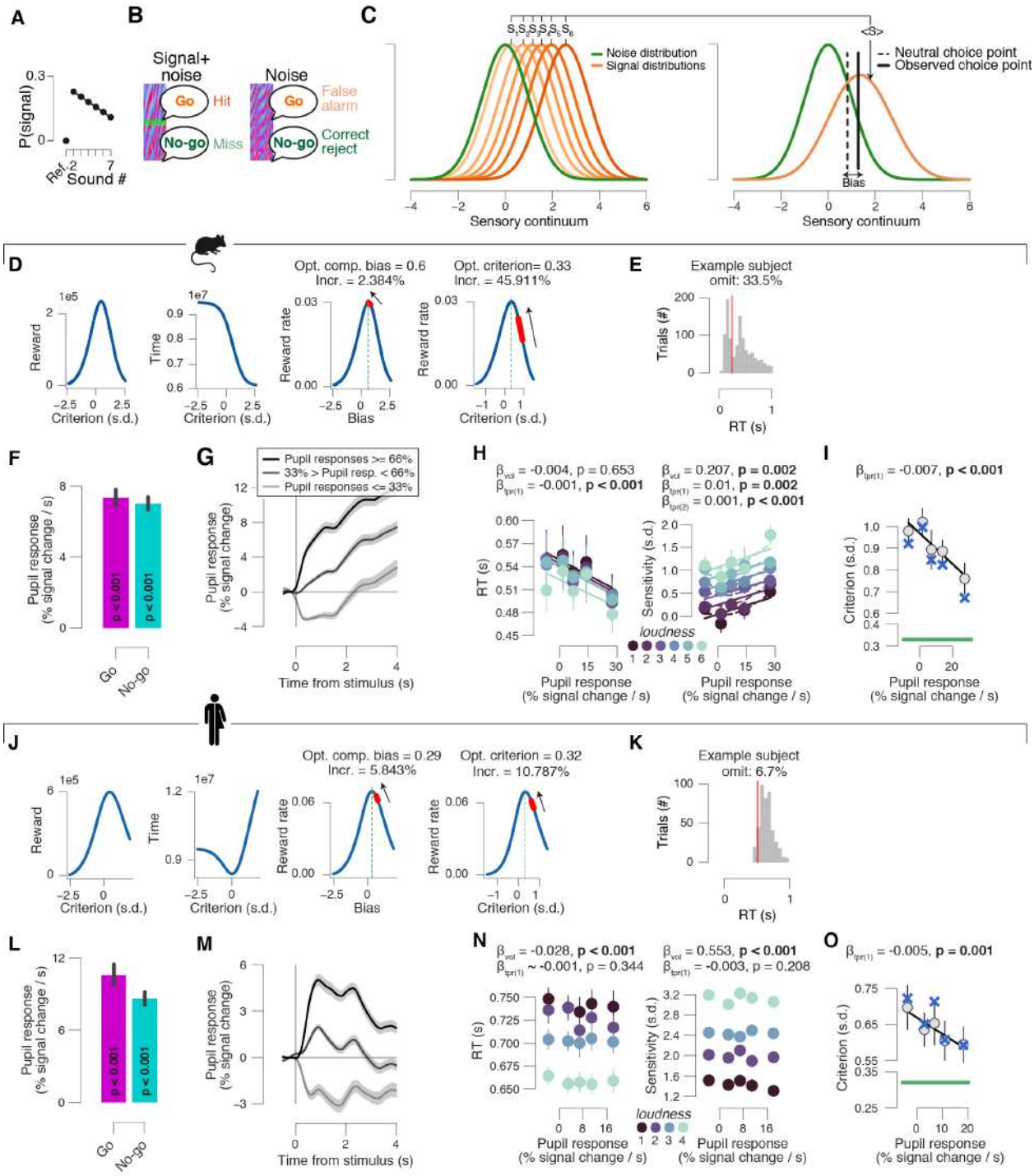


Figure S1: Quantifying pupil responses and behavior in mice and humans. (A) Probability of target signal across the sequence of 1–7 sounds. Hazard rate of signal occurrence was kept approximately flat. **(B)** The four combinations of stimulus category (signal+noise vs. noise) and behavioral choice (go 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). **(H)** 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. **(I)** As H, 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.

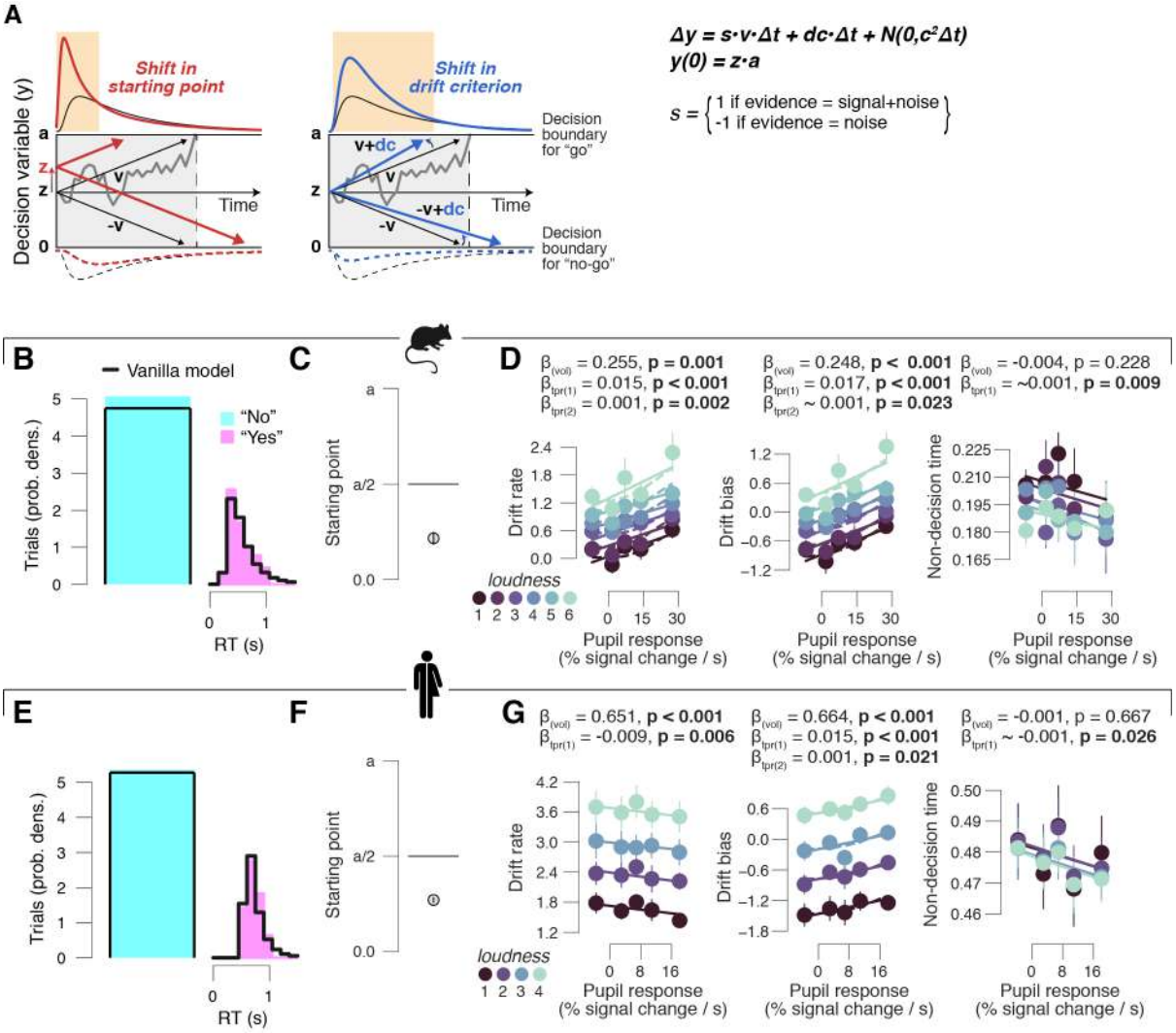


Figure S2: Pupil-dependent changes in computational model parameters during go/no-go task. (A) Schematic of drift diffusion model accounting for choices, and their associated RTs (for go-trials). Orange windows, RTs for which biased choices are expected under shifts in either "starting point" (z; left) 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.

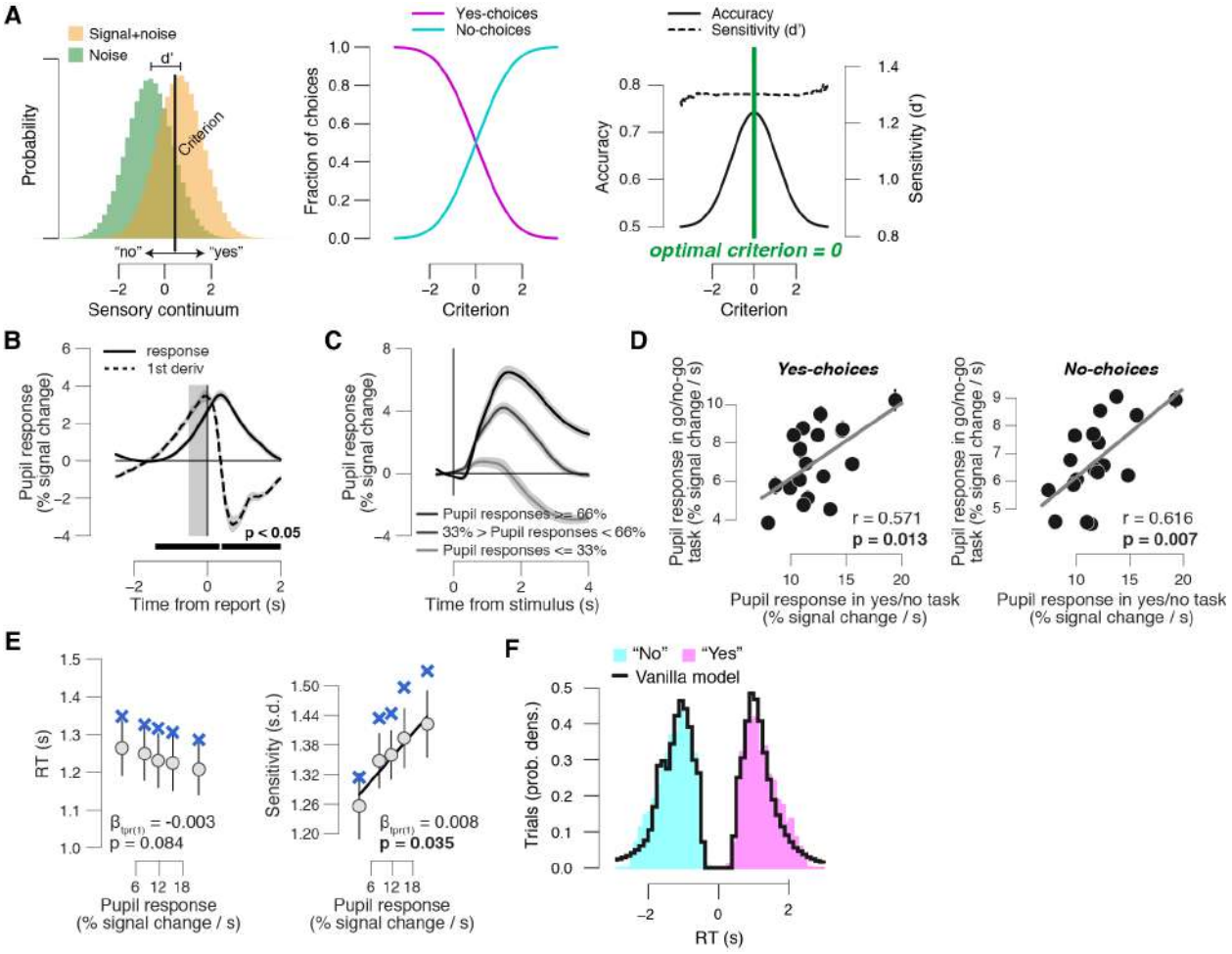


Figure S3: Optimality analysis of yes/no task. (A) Simulation study of optimal bias in yes/no (forced-choice) task (Materials and Methods). Optimality was defined as the level of choice bias (signal detection criterion) that maximized accuracy. Left: simulated internal sensory representations for signal+noise 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.

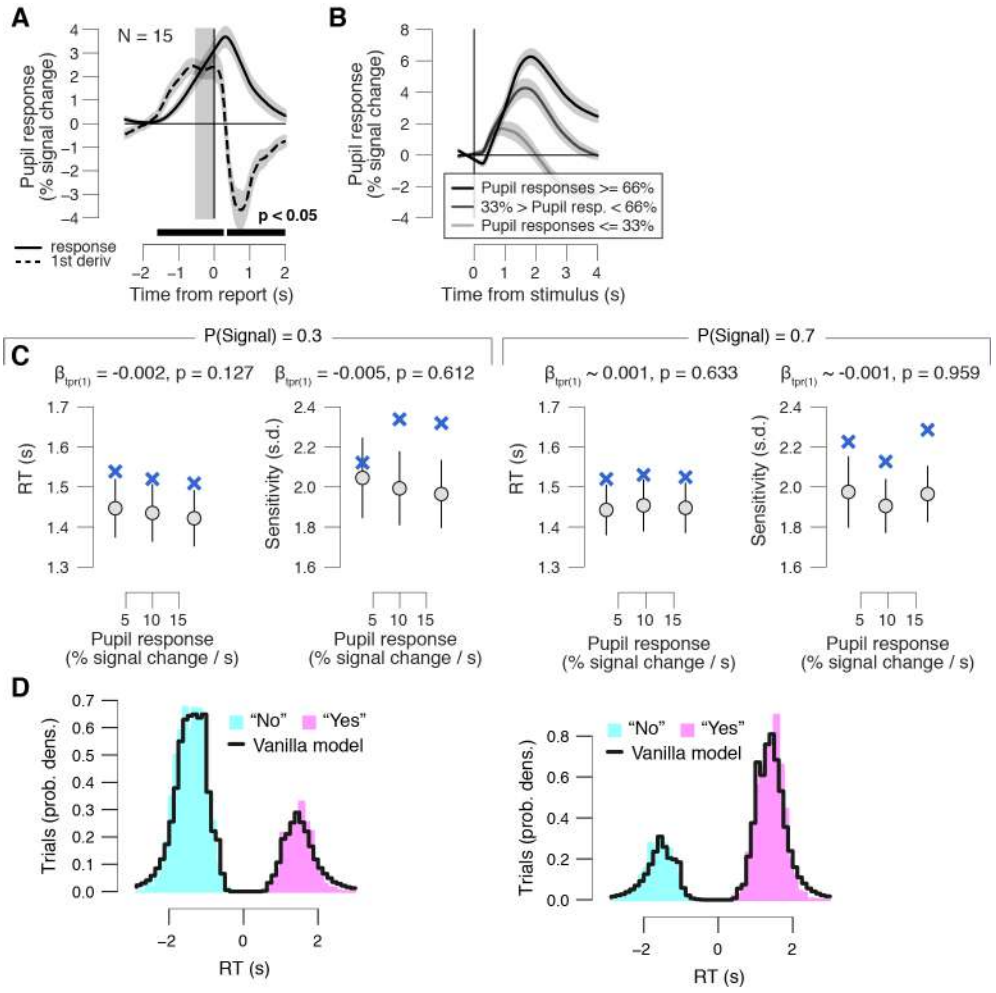


Figure S4: (A) Task-evoked pupil response (solid line) and response derivative (dashed line). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. Stats, paired-samples t-test. **(B)** Overall pupil response time course for three pupil 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).

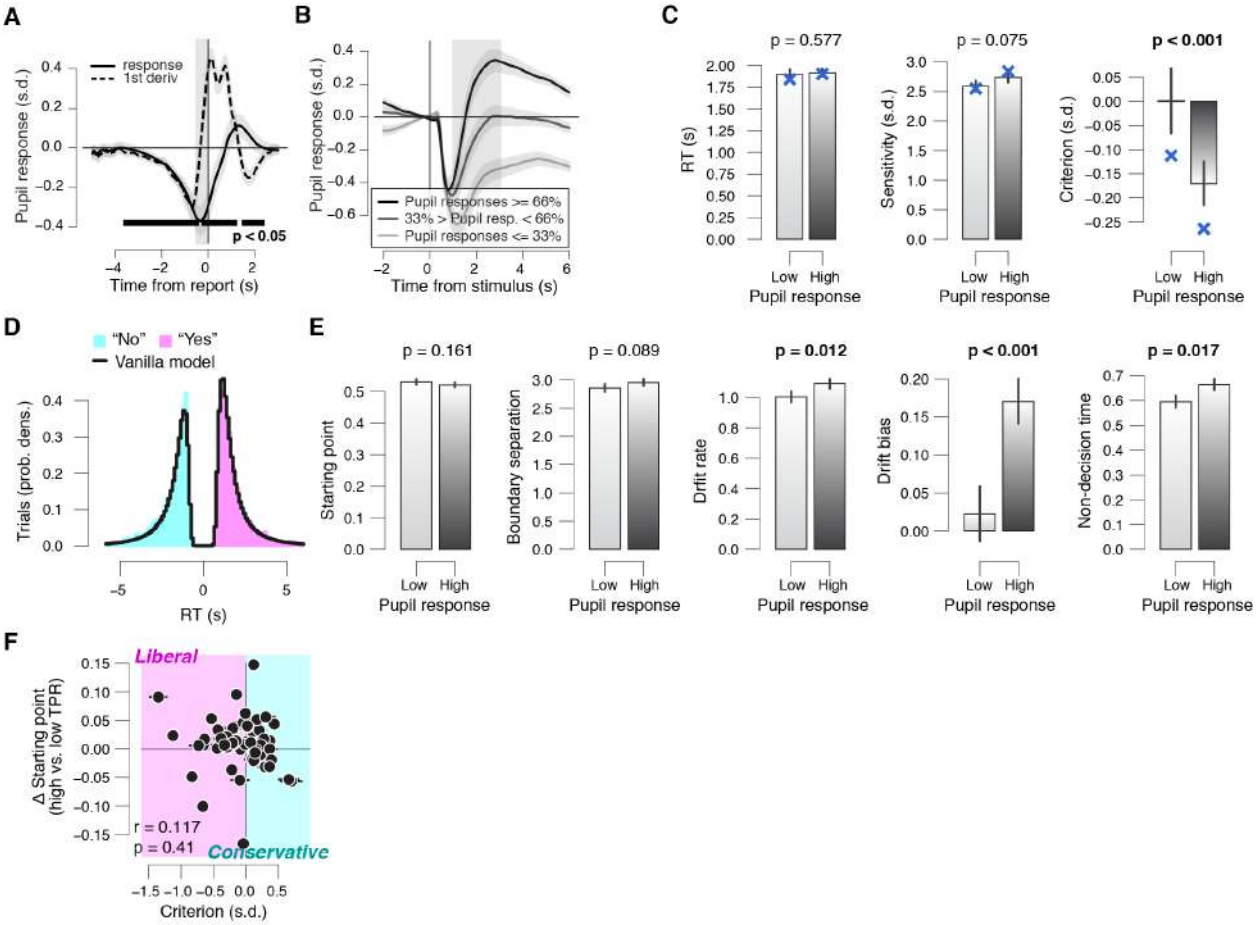


Figure S5: (A) Task-evoked pupil response (solid line) and response derivative (dashed line). Grey, interval for task-evoked pupil response measures (Materials and Methods); black bar, significant pupil derivative. Stats, paired-samples t-test. **(B)** Overall pupil response time course for three pupil 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).

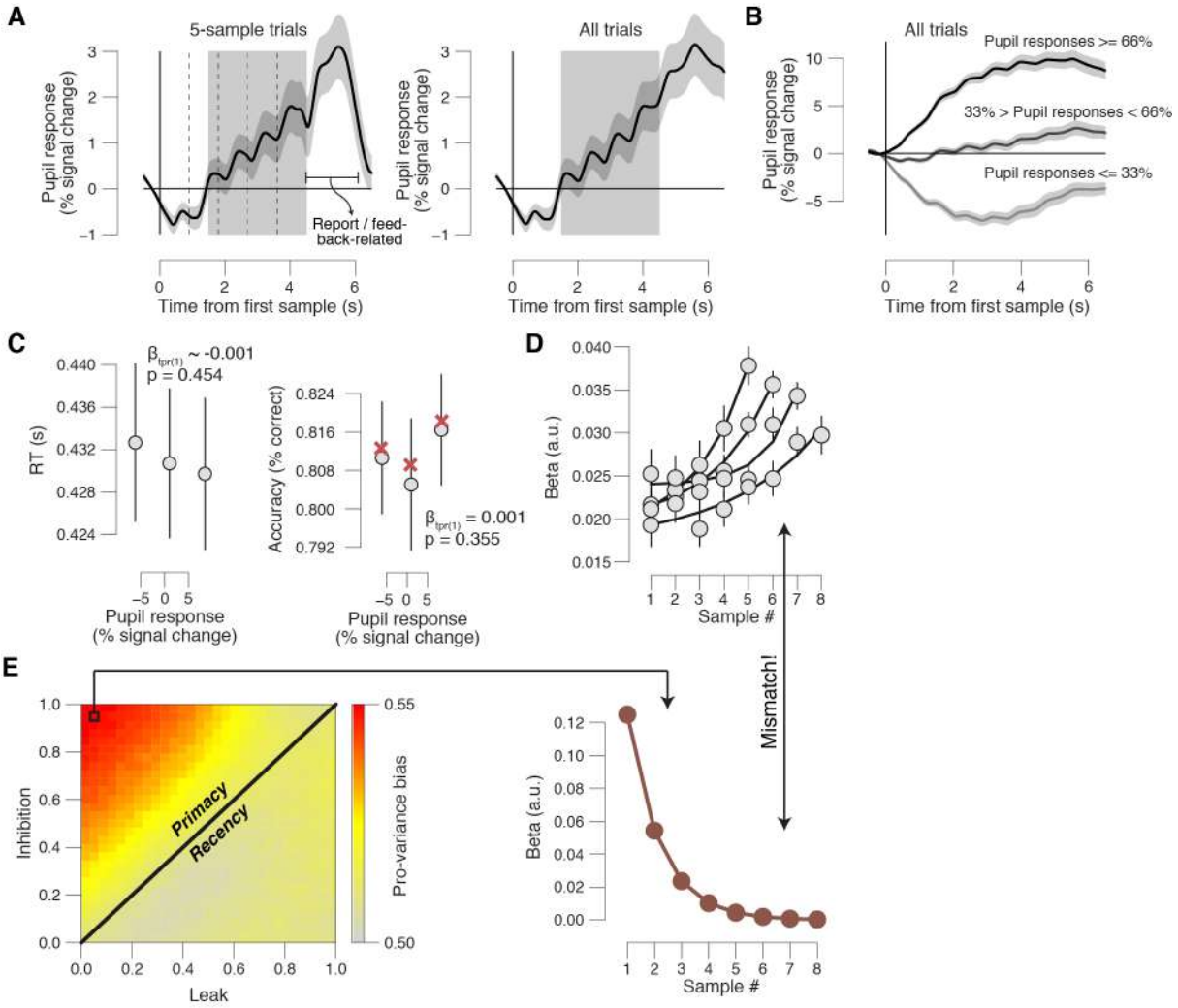


Figure S6: Pupil responses and pupil-dependent changes in behavior during value-based choice task. **(A)** Task-evoked pupil response time courses locked to the onset of the first pair of samples. Left: for the shortest trials (5 pairs of samples). Dashed vertical lines, onset of sample pairs 2-5. Right: 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.

Status: in preparation
de Gee JW, Wilming N & Donner TH (2019).

3

Phasic arousal shapes evidence accumulation by tuning cortical excitability

Conceptualization by JWG, THD; Investigation by JWG; Formal analysis by JWG, NW; Writing–original draft by JWG.

Many decisions are made in the face of uncertainty and involve accumulating ambiguous evidence over time. The brain's arousal systems are rapidly activated during such decisions. Pupil responses (a proxy of phasic arousal) predict a suppression of suboptimal biases in evidence accumulation across a wide range of decision-types and species. What are the biological underpinnings of this relationship? With simultaneous pupillometry and MEG we identified choice-predictive and pupil-linked oscillatory power in visual regions, established that their spectral profiles were highly similar, driven by common effects in the alpha-band (8–12 Hz) range. Critically, we the arousal-predicted suppression of evidence accumulation bias was formally mediated by the arousal-related shaping of choice-predictive alpha-band activity in regions along the dorsal stream of the visual cortical hierarchy (including V1) and in regions of the medial frontal cortex. Taken together, our results suggest that phasic arousal shapes evidence accumulation by tuning cortical excitability.

3.1 Introduction

The global arousal state of the brain changes continuously throughout the day [6, 11]. Changes in global arousal state are controlled in large part by subcortical nuclei, which, via widespread neuromodulatory

projections, can profoundly change the operating mode of target cortical circuits [6–10]. These same arousal systems are phasically recruited during decisions [6, 29–31, 45, 74, 125, 125–127].

The rapid arousal modulations during decisions are functionally relevant. Most decisions – including judgments about weak sensory patterns embedded in time-varying noise – are based on a protracted deliberation process that involves accumulating pieces of evidence in favor of a particular choice [20]. Using pupil dilation as a proxy of cortical arousal state [76, 78, 79], it was recently shown that phasic arousal is elevated already during such protracted deliberation processes [40, 42]. Furthermore, the functional impact of arousal transients during decision-making, which generalizes across species (humans and mice) and behavioral tasks (from perceptual to memory-based and numerical decisions), seems to be to suppress maladaptive biases in the accumulation of evidence leading up to a choice [33, 44]. These results were obtained by combining linear systems analysis to pupil data with rigorous modeling of behavior; the biological underpinnings of these effects remain however unknown.

Protracted decision-making is implemented by a distributed brain network: in the context of a perceptual decision, sensory cortex encodes input signals ("evidence samples"), which association cortex accumulates over time into a decision variable, and motor regions translate the decision into a behavioral act [18, 20, 22, 23]. Here we asked when and where in the cortical decision-making machinery phasic arousal shapes decision computations.

The modeling results described above put strong constraints on the expected shaping of cortical decision processing by phasic arousal. Specifically, the selective effect of phasic arousal on drift bias (rather than the over efficiency of evidence accumulation) predicts shaping of choice-related signals, and not of stimulus-related signals. Choice-predictive signals exist already in the first stages of visual processing [54, 166]. Indeed, in visual regions, pre-stimulus oscillatory activity in the alpha-band (about 8–12 Hz) has been shown to correlate with decision bias [167–170]. Intriguingly, phasic pupil dilation predicts oscillatory power in visual regions in the same frequency range [171]. We thus predicted that, in the context of a visual decision-making task, phasic arousal exerts its effect on evidence accumulation and choice by shaping oscillatory alpha-band activity in visual regions, perhaps as early as primary visual cortex (V1).

We tested this prediction in human subjects performing a simple contrast-detection task during simultaneous pupillometry and MEG recordings. Highlighting the quality of our data, we first established the common behavioral correlate of pupil-linked phasic arousal (reduction of drift bias) and a number of common MEG signatures of perceptual decision-making. We next identified choice-predictive and pupil-linked oscillatory power in visual regions and established that their spectral profiles were highly similar, driven by common effects in the alpha-band (8–12 Hz) range. Finally, we observed that the effect of phasic arousal on drift bias was formally mediated by the arousal-related shaping of alpha-band activity in regions along the dorsal stream of the visual cortical hierarchy (including V1) and in regions of the medial frontal cortex.

3.2 Results

During simultaneous pupillometry and MEG recordings, subjects performed a yes/no contrast detection task that on each trial involved discriminating signal+noise from noise stimuli, confidence ratings and feedback (Fig. 1A; Materials and Methods). We designed two conditions to experimentally manipulate choice bias (Materials and Methods). In the "neutral condition" the dynamic noise refreshed at 24 Hz and in the "conservative condition" the dynamic noise refreshed at 6 Hz. The two conditions were randomly interleaved across trials. Two independent online Quest staircases [172] kept subject's performance at roughly 75% correct in each bias condition (Materials and Methods).

3.2.1 Phasic arousal predicts a reduction of perceptual choice bias

We first aimed to establish the common behavioral correlate of pupil-linked phasic arousal: a reduction of drift bias [33, 44] (see also [74, 125, 126, 173]). To track phasic arousal, we measured the maximum of the rising slope of the pupil [39, 44, 174] in the second half of the stimulus interval (Fig. 1B). We chose this measure for two reasons: (i) to eliminate contamination by contrast-evoked constrictions and button-press related dilations (Materials and Methods), and (ii) to most specifically track noradrenergic activity [39]. From the resulting task-evoked pupil response measures we removed, via linear regression, components explained by signal contrast and baseline pupil size (Materials and Methods).

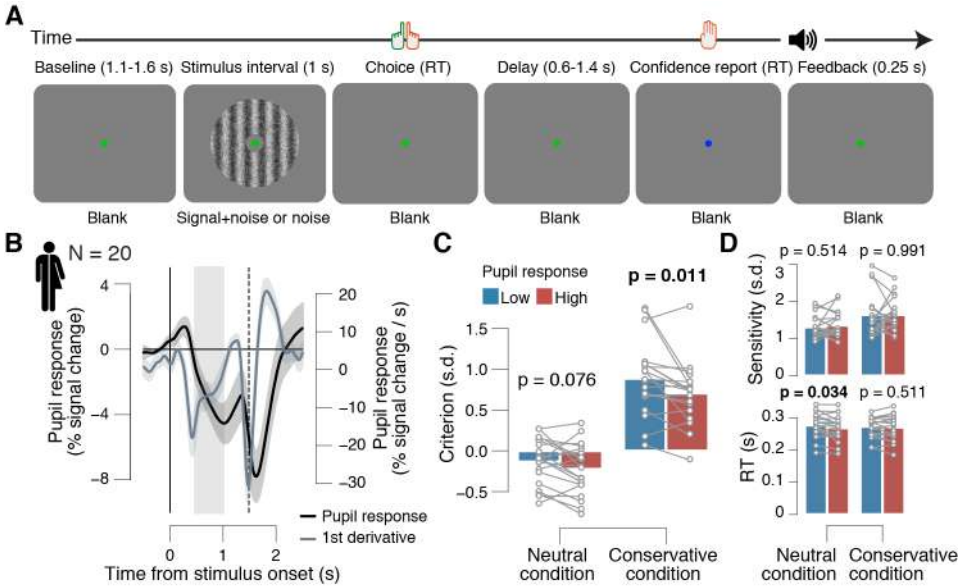


Figure 1: High phasic arousal is associated with reduced perceptual choice bias. (A) Visual contrast detection task. Schematic sequence of stimuli during a trial. Subjects were instructed to respond to a weak signal (gabor patch) in noise. In two conditions that were interleaved across trials, the noise refreshed at 24 Hz or 6 Hz (Materials and Methods). **(B)** Task-evoked pupil response (black) and response derivative (grey). Grey window, interval for task-evoked pupil response measures (Materials and Methods). Shading, s.e.m. **(C)** Relationship between overall perceptual choice bias (Materials and Methods) and task-evoked pupil response). Data points, individual subjects; stats, paired-samples t-test. **(D)** As panel C, but for perceptual sensitivity (d'; Materials and Methods; top) and reaction time (RT; bottom). Panels B-D: group average (N = 20).

We have previously shown that the effect of phasic arousal is adaptive: it predicts a suppression of bias of any sign (conservative and liberal) towards an optimal level [44]. In both conditions of the current experiment, signal+noise stimuli occurred on 50% of trials (Materials and Methods), and accuracy could be maximized by adapting a neutral bias. If the effect of phasic arousal is adaptive, it should predict no change in bias in the neutral condition (because subject's bias was already close to optimal), and a reduction of bias in the conservative condition (because subject's bias was maladaptive). This is exactly what we found (Fig. 1C). In the neutral condition, subjects had an overall weak liberal bias which was almost neutral (signal detection theoretic criterion c = -0.186, ± 0.060 s.e.m.), and there was no relationship between pupil-linked phasic arousal and bias (Fig. 1C). In the conservative condition, we found that subjects had an overall conservative perceptual choice bias (c = 0.725, ± 0.072 s.e.m.), a tendency to report "no", irrespective of the object evidence (Fig. 1C). This bias was suppressed on trials with large pupil responses (Fig. 1C). Phasic pupil responses exhibited no relationship to perceptual sensitivity (signal detection theoretic d') in either of the conditions (Fig. 1D) and predicted a (small) reduction of reaction time (RT) in the neutral condition (Fig. 1D).

3.2.2 Phasic arousal predicts a reduction of evidence accumulation bias

We have recently shown that the (monotonic) pupil-linked bias suppression is explained by a (monotonic) reduction of bias in the accumulation of evidence leading up to a choice [33, 44]. This mechanistic insight was achieved by fitting the drift diffusion model (Fig. S2A) to behavioral data. The drift diffusion model belongs to a class of bounded accumulation models of decision-making [2, 18, 19, 25] that describe the accumulation of noisy sensory evidence in a decision variable that drifts to one of two bounds. The model accounts well for behavioral data from a wide range of two-choice tasks [25].

Here, we again found evidence supporting the idea that phasic arousal interacts with the evidence accumulation process (Fig. 2). We used the hierarchical drift diffusion regression model to assess the trial-by-trial, linear relationship between the drift and pupil responses (Materials and Methods). Specifically, separately in each condition, we modeled the single-trial drift as a function of an intercept (capturing the overall "drift bias" towards "yes" or "no"), the stimulus (capturing the overall "drift rate", that is, the efficiency of evidence accumulation) and pupil response (capturing the trial-by-trial directed shifts in the drift [coded towards "yes"] as a function of phasic arousal) (see the equation in Fig. 2). The parameters starting point, boundary separation, non-decision time and drift rate variability were also included in the model, but varied only with condition, and not with phasic arousal (Fig. S2B,C). The evidence accumulation process was overall biased towards "no", especially in the conservative condition (Fig. 2A). Importantly, only in the conservative condition, this drift bias towards "no" was suppressed on trials characterized by large pupil responses (Fig. 2C).

We verified that (i) phasic arousal had no effect on the other model parameters, and (ii) the effect of phasic arousal on the drift was linear. To that end, we again fitted the drift diffusion model, but now separately per condition and separately for three pupil-response defined bins (Fig. S2D,E; Materials and

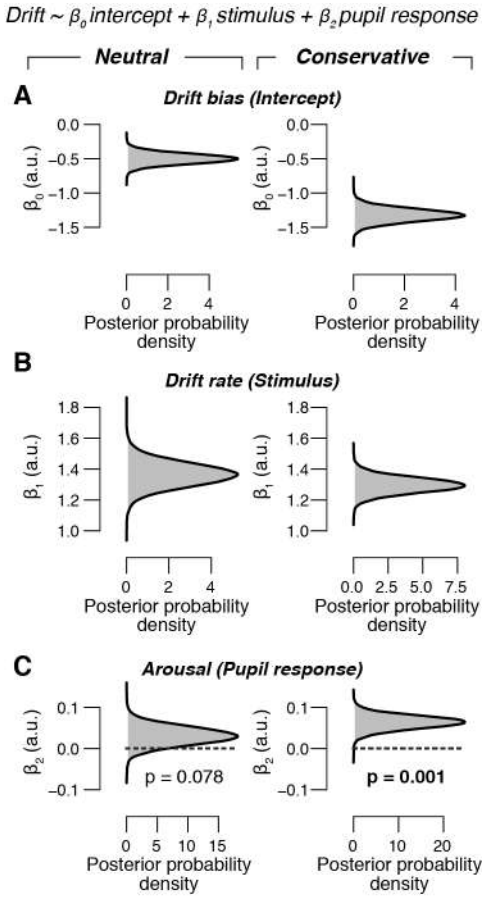


Figure 2: High phasic arousal is associated with reduced evidence accumulation bias. Group-level posterior probability densities in the hierarchical drift diffusion regression model for means of **(A)** intercept (capturing the overall drift bias), **(B)** stimulus (capturing the overall drift rate) and **(C)** pupil response (capturing the arousal-predicted directed shift in drift towards "yes"). The p-value quantifies the fraction of MCMC samples being smaller than 0 (dashed line) (Materials and Methods). All panels: group average (N = 20). See Figure S2B for the remaining parameters "starting point", "boundary separation" and "non-decision time".

Methods). First, when comparing the lowest and highest pupil bin (containing the lowest and highest 33% of trials; Materials and Methods) we observed only a significant effect of pupil-linked phasic arousal on drift bias in the conservative condition (Fig. S2D,E), and not on any of the other parameters in either condition. Second, drift bias estimates for the medium pupil bin fell in between estimates for the low and high pupil bins (Fig. S2D,E).

In sum, phasic arousal suppressed bias in the accumulation of evidence leading up to a choice. We next used the MEG data to characterize the biological underpinning of this interaction. Our approach was as follows: (i) establish a number of common MEG signatures of perceptual decision-making, (ii) identify choice-predictive and pupil-linked oscillatory power in visual regions and quantify the similarity of their spectral profiles, and (iii) test whether the effect of phasic arousal on drift bias is formally mediated by such choice-predictive and pupil-linked oscillatory activity.

3.2.3 MEG signatures of perceptual decision-making

We first confirmed that our data in both experimental conditions contained the typical MEG signatures of (i) stimulus processing and (ii) response preparation [22]. Computing time-frequency power at each vertex location and then averaging across all vertices within a region of interest (ROI; Materials and Meth-

ods) allowed us to track MEG signatures of decision-making at full time-frequency resolution throughout the distributed decision-making network. The signal+noise or noise stimuli induced robust population responses with a characteristic dynamic signature [27, 28]: in visual regions, the MEG power increased in a broad high-frequency range (about 50–150 Hz), accompanied by a decrease in the low-frequency range below 40 Hz (Figure 3A). In motor and frontal regions, we observed only task-induced decreases in the low-frequency range below 50 Hz (Figure 3A).

We exploited the contralateral bias of the cortical motor system (i.e., stronger activity in the hemisphere contralateral to the motor movement) to identify the characteristic dynamic signature of response preparation [26, 85]: in a number of lateral frontal and motor regions, including the hand area of M1, gamma-band activity (about 60–100 Hz) was selectively enhanced contralateral to the button-press, and alpha- and beta-band activity (about 8–36 Hz) was selectively suppressed (Figure 3B). We observed the choice-predictive lateralization of low-frequency power more than 0.5 s before report (Figure 3B).

Since subject's performance was held at roughly 75% correct (Materials and Methods), stimulus identity (signal+noise vs noise) was highly predictive of choice ("yes" vs. "no"). Thus, the choice contrast TFRs presented in Fig. 3B could, at least to an extent, reflect the stimulus contrast. We ruled out this concern by recomputing the choice contrast as follows: (hit + false alarm) – (correct reject + miss), effectively

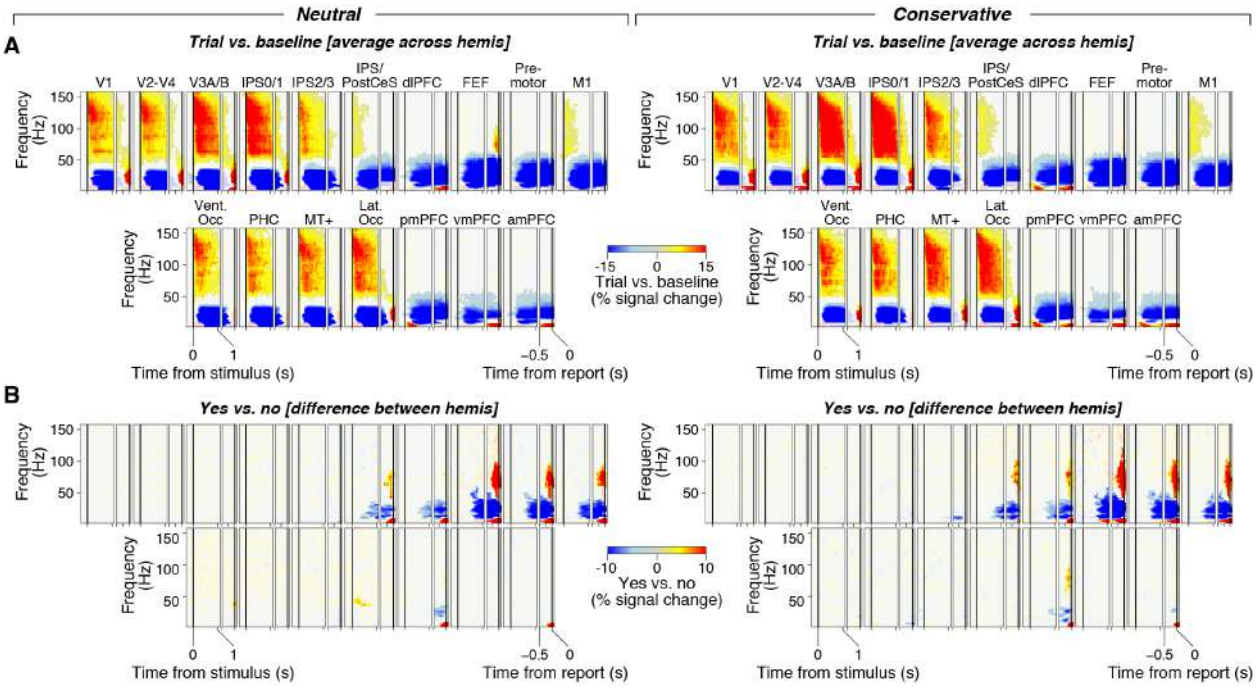


Figure 3: Time-frequency responses in the yes/no contrast detection task. (A) Overall task-induced time-frequency responses (average across hemispheres) in the selected ROIs, expressed as % signal change from baseline (Materials and Methods), separately for the neutral (left) and conservative condition (right). See Materials and Methods for the definition of each cortical ROI. **(B)** As panel A, but for the contrast "yes" vs. "no" and after subtracting data from the ipsi-lateral hemisphere from the contralateral hemisphere (Materials and Methods). See Fig. S3 for the choice contrast computed as (hit + false alarm) – (correct reject + miss), effectively factoring out effect of stimulus identity and accuracy. All panels: Group average (N = 20), saturation, significant time-frequency tiles ($p < 0.05$, two-sided permutation across subjects, cluster corrected).

factoring out effect of stimulus identity and accuracy. Doing so verified that the results in Fig. 3B indeed reflect pure choice-predictive activity (Fig. S3), consistent with earlier reports [26, 85].

In sum, in both experimental conditions we observed the typical MEG signatures of stimulus processing and response preparation, highlighting the quality of our data. We hypothesized that the arousal-related modulation of choice-predictive signals in visual cortex might underlie the arousal-related suppression of choice bias. Our aim was thus to first identify choice-predictive and pupil-linked oscillatory power in visual regions and quantify the similarity of their spectral profiles.

3.2.4 Similar spectral profiles of choice-predictive and pupil-linked oscillatory power in visual cortex

We recomputed the "yes" vs. "no" contrast, but now after averaging the activity of both hemispheres (Fig. 4A), instead of taking the difference between hemispheres (as in Fig. 3B). In visual regions, and in both experimental conditions, we observed a suppression of alpha-band activity (about 8–12 Hz) for "yes" vs "no" choices (Fig. 4A). This spectral profile closely resembled the pre-stimulus alpha-band activity that was previously shown to correlate with decision bias [167–170]. However, its temporal profile was distinct: rather than spontaneously fluctuating in between trials, it was closely time-locked to the behavioral report (button press) (Fig. 4A). As before with the laterization measures, we recomputed the choice contrast as follows: (hit + false alarm) – (correct reject + miss), effectively factoring out effect of stimulus identity and accuracy. Doing so verified that the results in Fig. 4A indeed reflect pure choice-predictive activity (Fig. S4).

The magnitude of task-evoked pupil responses predicted a suppression of alpha-band activity (about 8–12 Hz) in visual regions, which was closely time-locked to stimulus onset (Fig. 4B). The early (w.r.t stimulus-onset) pupil-linked alpha-band suppression was in line with the idea that phasic arousal is elevated already during protracted decisions [40, 175], and shapes the evidence accumulation process (Fig. 2; [33, 44]). Note that the significant timepoints before stimulus onset are due the window length of 400 ms in our time-frequency analysis (Materials and Methods) which caused some smearing in the low frequencies. A number of further observations indicated that the pupil-linked alpha-band suppression was task-induced: (i) we removed, via linear regression, components from the trial-wise pupil response measures that were explained by pre-trial baseline pupil size (Materials and Methods); thus, the pupil-linked alpha-band suppression was not a carry-over effect from the pupil response dependence on its baseline [40, 72]; (ii) the pupil-linked alpha-band suppression was also not due to changes in retinal illumination as a function of pupil dilation: it occurred already before the sluggish pupil response [40, 41] was different between the high and low pupil response bins (Fig. S1B); (iii) the pupil-linked alpha-band suppression identified here closely resembled the pupil-linked alpha-band suppression time-locked to perceptual switches in the context of a bistable stimulus [171].

To quantify the similarity between choice-predictive and pupil-linked oscillatory power, we correlated their spectral profiles (Fig. 4C,D). Specifically, we first computed the spectral profile of the choice con-

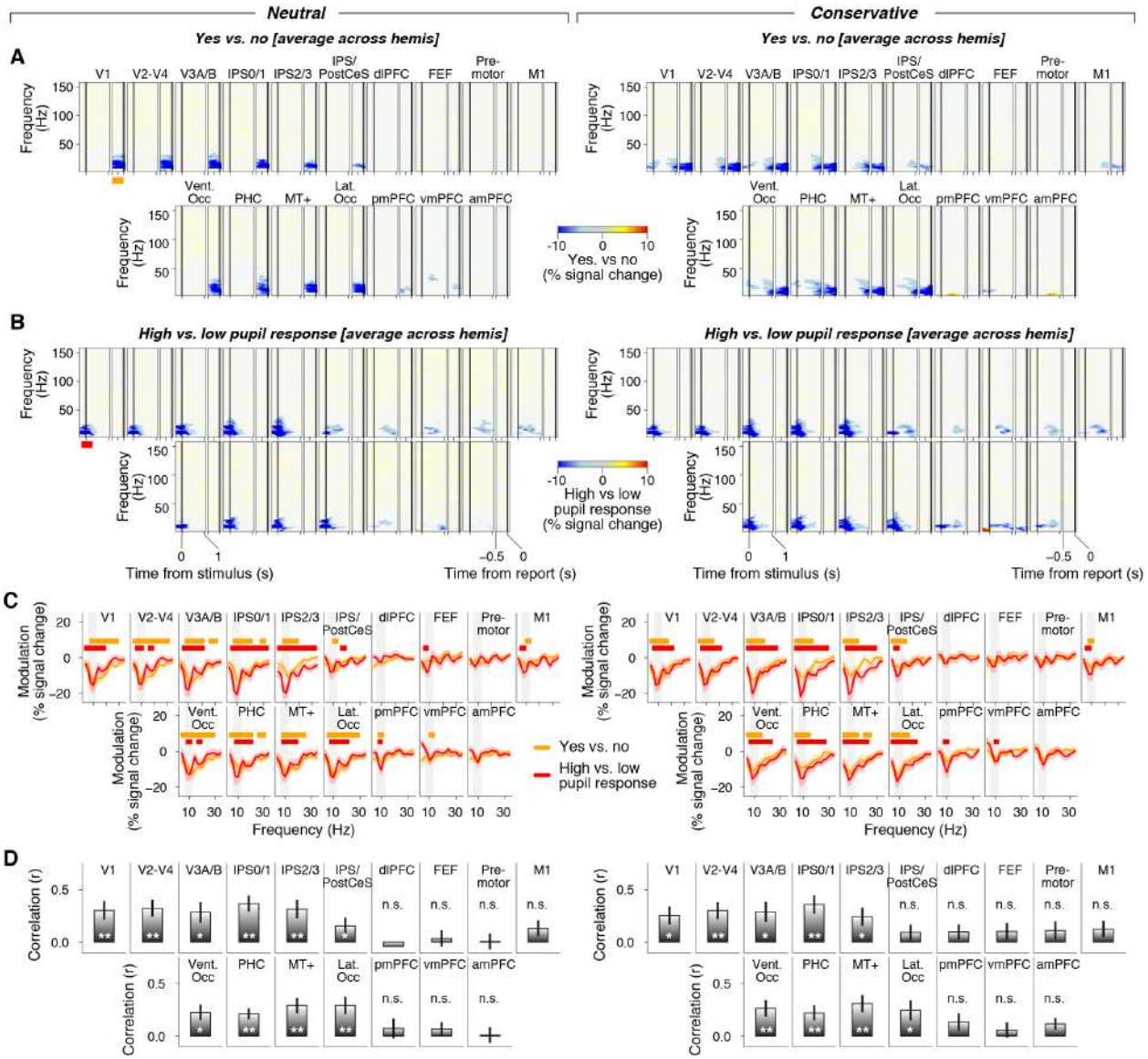


Figure 4: Overlapping spectral profiles for choice and pupil response contrasts. (A) "Yes" vs. "no" contrast of task-induced time-frequency responses (average across hemispheres) in the selected regions of interest (ROIs), expressed as % signal change from baseline (Materials and Methods), separately for the neutral (left) and conservative condition (right). See Fig. S4 for the choice contrast computed as (hit + false alarm) – (correct reject + miss), effectively factoring out effect of stimulus identity and accuracy. See Materials and Methods for the definition of cortical ROIs. Saturation, significant time-frequency tiles ($p < 0.05$, two-sided permutation across subjects, cluster corrected); orange bar (far left), timepoints used for calculating power spectra in panel C. **(B)** As panel A, but for the contrast high vs low pupil response (average across hemispheres). Note that the significant timepoints before stimulus onset are due the window length of 400 ms in our time-frequency analysis (Materials and Methods) which causes some smearing in the low frequencies. Red bar (far left), timepoints used for calculating power spectra in panel C. **(C)** Spectra of the power modulation of the "yes" vs. "no" contrast (average across hemispheres; time window, 0.5–0 s before report; see also orange bar in panel A) and high vs. low pupil response (average across hemispheres; time window, -0.2 s before stimulus onset to 0.3 s after; see also red bar in panel B), separately for the neutral (left) and conservative condition (right). Spectra are shown for frequencies 2–36 Hz; there were no significant power modulations in the higher frequencies 36-160 Hz (Fig. 3A,B). Bars, clusters of significant power modulation ($p < 0.05$, two-sided permutation across subjects, cluster corrected). **(D)** Correlation between the choice and pupil contrast spectra shown in panel C. Stats, Wilcoxon signed-rank test; *, $p<0.05$; **, $p<0.01$. All panels: group average (N = 20); Shading or error bars, s.e.m.

trast after collapsing across timepoints -0.5 to 0 s from report, and that of the pupil response contrast after collapsing across timepoints -0.2 to 0.3 s from stimulus onset (Fig. 4C). We observed significant correlations between choice-related and pupil-related spectra in each region of the visual cortical hierarchy, ranging from primary visual cortex all the way up to intraparietal regions (Fig. 4D).

In sum, in visual regions, phasic arousal predicted alpha-band suppression around stimulus onset time that closely resembled choice-predictive alpha-band suppression in the inter-trial interval [167–170] or just before report (Fig. 4D). This pattern of results is consistent with a scenario in which the effect of phasic arousal on drift bias (Figure 2) is mediated by the arousal-related shaping of alpha-band activity. We next used structural equation modeling (Materials and Methods) to formally test this idea.

3.2.5 The arousal-related shaping of visual alpha-band activity mediates the arousal-related shaping of evidence accumulation

The model is illustrated in Fig. 5A (Materials and Methods). We used hierarchical linear regression analysis to quantify the relationship between trial-wise pupil responses and trial-wise alpha-band activity measures in primary visual cortex in the time window -0.2 to 0.3 s from stimulus onset (the a-path; Materials and Methods). Replicating the effects reported in Fig. 4B, we observed a negative single-trial correlation between pupil responses and alpha-band activity in V1 (Fig. 5B and Fig. S5A,B left panels). We used the hierarchical drift diffusion regression model to quantify the relationship between trial-wise pupil responses, alpha-band activity in V1 and drift bias (the b-path and c'-path, respectively; Materials and Methods). We found that trial-wise alpha-band activity in V1 was negatively correlated to drift bias (towards "yes"), but only in the conservative condition (Fig. 5C and Fig. S5A,B right panels). Note that the single-trial drift was estimated now as a function of both alpha-band activity in V1 and pupil responses, which means that the observed relationship between alpha-band activity in V1 and drift bias existed over and beyond what could be explained by phasic arousal.

In the conservative condition, the relationship between phasic arousal and drift bias was indeed mediated by the arousal-related shaping of alpha-band activity in V1: the indirect path was highly significant ($p < 0.001$) (Fig. 5D). It was a partial mediation effect: the relationship between pupil responses and drift bias (Fig. 2) was not fully explained away by visual alpha-band activity (Fig. S5C,D).

We conducted a model-free and non-hierarchical version of the mediation analysis (Fig. S6A; Materials and Methods) in order to characterize the spatial topography of these relationships throughout the visual cortical hierarchy and beyond. Replicating the result from the hierarchal approach described above (Fig. 5), in the conservative condition, we again found evidence for a mediation by the arousal-related shaping of alpha-band activity in V1 (Fig. S6B-D). Furthermore, in the conservative condition, we observed similar effects in the dorsal stream of the visual cortical hierarchy (specifically in the clusters V2-4, V3A/B, IPS2/3 and IPS/PostCeS), in medial frontal regions, and in the hand area of M1 (Fig. S6B-D). We did not find evidence for mediation by alpha-band activity in the ventral stream of the visual cortical hierarchy or in lateral frontal regions (Fig. S6B-D).

Because the physical stimulus was included as a covariate in the all mediation models, our results imply that the interplay between phasic arousal and cortical alpha-band activity accounted for a significant component of behavioral variability, over and above the objective sensory evidence.

3.3 Discussion

Using pupil dilation as a proxy of cortical arousal state [76, 78, 79], it was recently shown that phasic arousal is elevated already during protracted deliberation processes [40, 175], and suppresses maladaptive biases in the accumulation of evidence leading up to a choice [33, 44]. Providing direct support to these insights, we (i) replicated the behavioral correlate of phasic arousal (reduction of drift bias), (ii) identified choice-predictive and pupil-linked oscillatory power in visual regions, (iii) established that their spectral

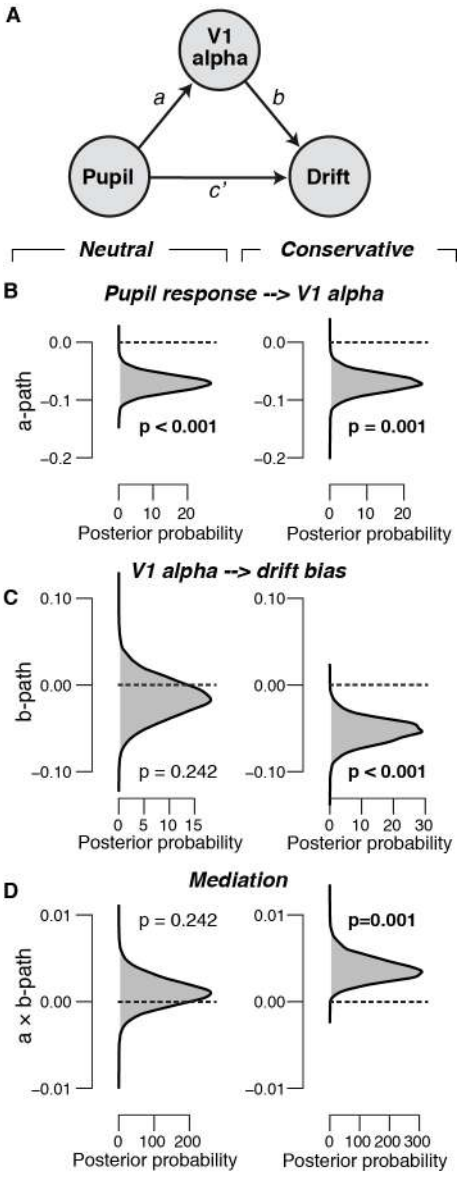


Figure 5: Relationship between pupil-linked arousal, alpha-band activity in V1 and drift (A) Schematic of mediation analysis (see Materials and Methods). Pupil, vector of trial-wise pupil responses; V1 alpha, vector of trial-wise alpha-band activity measures in primary visual cortex (Materials and Methods); drift, single-trial drift towards the "yes" bound, as estimated with the hierarchical drift diffusion regression model. All arrows are regressions. **(B)** Group-level posterior probability densities from hierarchical regression (Materials and Methods) of the linear relationship between trial-wise pupil responses and V1 alpha. The p-value quantifies the fraction of MCMC samples being larger than 0 (dashed line) (Materials and Methods). We verified that the effect of phasic arousal on alpha-band activity is linear (Fig. S5A,B left panels). **(C)** Group-level posterior probability densities from the hierarchical drift diffusion regression model of the linear relationship between trial-wise V1 alpha and drift bias. The p-value quantifies the fraction of MCMC samples being larger than 0 (dashed line) (Materials and Methods). We verified that the effect of V1 alpha on drift bias is linear (Fig. S5A,B right panels). **(D)** Group-level posterior probability densities of the indirect path a × b (Materials and Methods). The p-value quantifies the fraction of MCMC samples being smaller than 0 (dashed line). See also Fig. S6 for mediation analysis in other cortical ROIs. All panels: Group average (N = 20); stats, directly on the posterior probability densities (Materials and Methods).

profiles were highly similar, driven by common effects in the alpha-band (8–12 Hz) range, and (iv) found that the effect of phasic arousal on drift bias was formally mediated by the arousal-related shaping of choice-predictive alpha-band activity in regions along the dorsal stream of the visual cortical hierarchy (including V1) and in regions of the medial frontal cortex.

We here used pupil responses as a peripheral readout of changes in cortical arousal state [11, 32]. Changes in pupil diameter are also associated with locus coeruleus (LC) responses in humans [33, 34], monkeys [35, 36], and mice [37–39]. 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 [33, 35, 39]. Thus, we remain agnostic about the exact neuroanatomical source(s) of the reported effects of phasic arousal.

One account holds that task-evoked LC responses during decision-making are triggered by a threshold crossing in some cortical circuit accumulating evidence for different choice options [6]. The resulting cortex-wide noradrenaline release then facilitates the translation of the choice into a motor act. This model has been supported by observations from monkeys performing simple perceptual decisions, that phasic responses of LC neurons are more closely aligned to behavioral responses than to the onset of stimuli [47, 48] and encode decisions to execute, but not withhold, movements [49]. Our results do not support this account. First, the physiological correlates of pupil-linked phasic arousal (alpha-band suppression in visual regions) occurred closely time-locked to stimulus onset, and not around the behavioral report (button press). Second, within the drift diffusion model, the "post-decisional" account [6] predicts a reduction in non-decision time and no effect on evidence accumulation. In contrast to this prediction, we found that phasic arousal suppressed drift bias, but did not predict changes in non-decision time. Together, our results imply 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.

How does the arousal-mediated suppression of alpha-band activity (increase of neural excitability) explain the reduction of drift bias in the conservative condition? Biophysically detailed models of the neural computations underlying decision-making [18] help to conceptualize this question. In one specific scheme for yes/no choices, supported by neurophysiological data [26, 50, 51], two neural populations accumulate evidence for "yes" and "no" toward separate bounds, and compete via mutual inhibition [26, 52, 53]. The yes-population accumulates the sensory evidence for signal presence–i.e., neural activity in visual cortex, which is either stimulus induced (on signal+noise trials) or spontaneous (on noise trials) [54]. The no-population accumulates a "default input" [52]. If, for example, visual cortex is larger than the region providing the "default input", then boosting overall neural excitability will asymmetrically enhance neural activity in visual cortex (on both signal+noise and noise trials), which, in turn, will translate in faster accumulation toward "yes".

Our results are remarkably consistent with two lines of research which were so far largely disconnected. First, our results support the notion that cortical alpha-band activity reflects excitability and predicts changes in choice bias rather than accuracy [167–170, 176]. Second, our results are in line with a body of work that shows that pupil-linked arousal boosts cortical excitability and suppresses low-frequency synchronous activity [76, 78, 79, 174], including alpha-band activity [171]. The release of noradrenaline and/or acetylcholine might explain this relationship: these modulatory neurotransmitters tune ionic currents, which not only control the membrane potential and excitability of cortical neurons, but also strongly suppress the generation of slow oscillations. Thus, the suppression of low frequency activity need not purely reflect a (attentional) top-down signal from higher-order sensory regions or from association cortices [177–179], but might also reflect the impact of neuromodulators released by subcortical nuclei.

Taken together, our results suggest that the phasic arousal shapes evidence accumulation by tuning cortical excitability. Thus, when decisions are made in the face of uncertainty, tracking phasic arousal signals may be just as important for predicting cortical decision-computations and subsequent choice behavior as tracking the objective evidence gathered from the outside world.

3.4 Materials and Methods

3.4.1 Subjects

Twenty-three subjects (15 female, age range 20-31) participated in the experiment, entailing concurrent pupillometry and MEG recordings. All had normal or corrected to normal vision and no history or indications of psychological or neurological disorders. The experiment was approved by the ethics committee of the University Medical Center Hamburg-Eppendorf. Subjects performed between 7 and 10 blocks of 128 trials each (distributed over two measurement sessions, each lasting for approximately 75 minutes), yielding a total of 896-1280 trials per participant. Subjects completed one training session on a separate day before the MEG recordings. We also collected a structural MRI for each subject on a separate day after the MEG recordings. Subjects were paid 10€ per hour. Three subjects completed only one MEG recording session and were excluded from further analyses.

3.4.2 Behavioral task

Each trial consisted of six consecutive intervals (Figure 1A): (i) the baseline interval (uniformly distributed between 1.1 and 1.6 s) containing only a green fixation dot; (ii) the stimulus interval (1 s duration) containing either a signal (gabor patch) in noise or only noise; (iii) the choice interval, the start of which was signaled by the disappearance of the stimulus, and terminated by the subject's response; (iv) delay interval (uniformly distributed between 0.6 and 1.4 s); (v) confidence interval, the start of which was signaled the fixation dot turning blue, and terminated by the subject's response; (vi) feedback interval (0.25 s).

A dynamic noise pattern was presented throughout the stimulus interval. The luminance across all pixels was kept constant. This pedestal noise pattern had 20% contrast. We designed two conditions to experimentally manipulate choice bias. In the "neutral condition" the dynamic noise refreshed at 24 Hz and in the "conservative condition" the dynamic noise refreshed at 6 Hz (every four frames). The two conditions were randomly interleaved across trials, under the constraint that they would occur on 50% of the trials within each block of 128 trials.

On one half of trials ("signal+noise" trials), a sinusoidal grating (2 cycles per degree) was superimposed on the visual noise for the entire stimulus interval (Figure 1A). The other half of trials ("noise" trials) contained no target signal during the stimulus interval. Signal presence was randomly selected on each trial, under the constraint that it would occur on 50% of the trials within the neutral and conservative conditions, and within each block of 128 trials. All signal and noise stimuli were presented in a Gaussian annulus, with an average distance (± SD) to fixation of 4.8 (1.8) degrees (Figure 1A).

After stimulus offset subjects were instructed to report the presence or absence of the signal ("interrogation paradigm") by pressing one of two response buttons with their left or right index finger. The mapping between perceptual choice and button press (e.g., "yes" –> press right key; "no" –> press left key) was counterbalanced across the two recording sessions. After a short delay, subjects were prompted to rate their confidence on a four-point scale ("not at all", "not so much", "quite", "very") with the four fingers of their right hand. Subjects received auditory feedback (high tone –> correct; low tone –> error) at the end of each trial.

Throughout the main experiment, the contrast of the target signal was dynamically titrated to the threshold level that yielded about 75% correct choices using an adaptive staircase procedure (Quest) [172]. We ran independent staircases for the neutral and conservative conditions. The mean threshold contrasts were 1.128 % (± 0.016 % s.e.m.) in the neutral condition and 1.221 % (± 0.009 % s.e.m.) in the conservative condition (Fig. S1A). These threshold contrasts yielded a mean accuracy of 73.36 % correct (± 0.43 % s.e.m.) in the neutral condition, and 71.60 % correct (± 0.44 % s.e.m.) in the conservative condition.

Stimuli were back-projected on a transparent screen using a Sanyo PCL-XP51 projector with a resolution of 800x600 at 24 Hz. The screen was positioned 65 cm away from their eyes. The luminance profile was linearized by measuring and correcting for the systems gamma curve. A doubling of contrast values therefore also produced a doubling of luminance differences. During the first training session stimuli were presented on a VIEWPixx monitor with the same resolution and refresh rate (also linearized). To minimize any effect of light on pupil diameter, the overall luminance of the screen was held constant throughout the experiment.

3.4.3 Data acquisition

We used a CTF MEG system with 275 axial gradiometer sensors and recorded at 1200 Hz. Recordings took place in a dimly lit magnetically shielded room. We concurrently collected pupillometric data at 1000 Hz using an EyeLink 1000 Long Range Mount (SR Research, Osgoode, Ontario, Canada). We continuously

monitored head position by using three fiducial coils. We used Ag/AgCl electrodes to measure ECG and vertical and horizontal EOG.

3.4.4 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) 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 [108], 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 We computed task-evoked pupil responses time-locked to the stimulus. We used stimulus-locking because the high contrast dynamic noise pattern, which occurred in all trials, elicited a transient pupil constriction (Fig. 1B). Thus, locking pupil responses to the stimulus balanced those pupil constrictions 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 [33, 39, 174] in the 500 ms before stimulus offset. This time window of interest did not include motor responses, so that button-press related dilations [40, 133] were also confounding factor. From the resulting task-evoked pupil response measures we removed, via linear regression, components explained by signal contrast and baseline pupil size (Materials and Methods), so that all the results reported here could not be due a carry-over effect from the pupil response dependence on its baseline [40, 72].

In some analyses (Figs. 1 and 4), trials were sorted by pupil response amplitude and collapsed into three bins containing the lowest and highest 33% (which were used for analyses), as well as the intermediate 33% of pupil response amplitudes. This achieved a trade-off between maximizing both (i) trial counts in the high and low pupil response bins and (ii) the disparity between the pupil responses for both bins. The resulting pupil bins were associated with different overall pupil response amplitudes across the whole duration of the trial (Fig. S1B). In other analyses (Figs. 2 and 5), we quantified the interaction between pupil responses, cortical responses and behavior at the single-trial level.

3.4.5 Analysis of task-induced MEG responses

Preprocessing We automatically labeled artifacts in raw MEG data by detecting blinks, muscle artifacts, sensor jumps and cars passing by in the vicinity of the building. Blinks were detected based on

concurrently recorded eye-movement signals (SR-Research EyeLink 1000). Sensor jumps were detected by convolving each sensor with a filter designed to detect large sudden jumps, and subsequently by looking for outliers in the filter response. Muscle (car) artifacts were detected by filtering each channel in the 110-140 Hz (<1 Hz) range and by detecting outliers that occurred simultaneously in many channels. To remove power line noise, we applied a notch filter (50 and its harmonics). In a next step we epoched data, downsampled to 600 Hz and discarded all epochs that contained artifacts. The code that detects artifacts and produces source estimates (see below) is available at `www.github.com/Donnerlab/pymeg`.

MEG SOURCE RECONSTRUCTION We used beamforming to reconstruct the sources of activity observed at the MEG sensor level, as follows. First, we constructed individual three-layer head models from subject specific MRI scans using the Fieldtrip functions ft_volumesegment and ft_prepare_mesh [180]. Head models were aligned to the MEG data by a transformation matrix that aligned the average fiducial coil position in the MEG data and the corresponding locations in each head model. We computed one transformation matrix per recording session. Second, we reconstructed cortical surfaces from individual MRIs using Freesurfer [116, 181] and aligned two different atlases to each surface [182, 183]. Third, we used MNE [184] to compute LCMV filters that project data into regions of interest defined by each atlas [185]. LCMV filters combined a forward model based on the head model and a source space constrained to the cortical sheet (4096 vertices per hemisphere, recursively subdivided octahedron) with a data covariance matrix estimated from the cleaned epoched data. We computed one filter per ROI, based on the covariance matrix computed on the time-points from stimulus onset until stimulus offset (1 s after stimulus onset) across all trials. We chose the source orientation with maximum output source power at each cortical location. In a final step, we computed a multi-taper based TFR decomposition of our epoched MEG data and projected the complex time series into source space. For low frequencies (2–40 Hz in steps of 2 Hz), we used the multitaper technique with five discrete proloid slepian tapers [186], a window length of 400 ms in steps of 60 ms, and 2.5 Hz frequency smoothing. For high frequencies (40-160 Hz in steps of 4 Hz), we used the multitaper technique with one proloid slepian taper, a window length of 400 ms in steps of 60 ms, and 7.5 Hz frequency smoothing. In source space we computed TFR power at each vertex location and then averaged across all vertices within a ROI. We aligned the polarity of time-series at neighboring vertices, because the beamformer output potentially includes arbitrary sign flips for different vertices.

We converted data in each region of interest (ROI) to units of modulation (percent signal change). Data for each ROI and frequency was normalized by the corresponding (ROI and frequency-wise) mean power across all trials in the time window 0.25 to 0.15 s before stimulus onset. For subsequent analyses we either averaged ROIs across hemispheres or computed a measure of lateralization. For the average measure we averaged both hemispheres per ROI. For the lateralization score we subtracted data from the ipsi-lateral hemisphere from the contralateral hemisphere.

Quantification of trial-wise alpha-band activity measures Separately per region of interest, we computed trial-wise alpha-band activity measures as the mean alpha power (8–12 Hz) in the time window -0.2–0.3 s from stimulus onset. The resulting single trial measures were normalized by the corresponding (ROI-wise) mean power in the same frequency range across all trials in the time window 0.25 to 0.15 s before stimulus onset.

3.4.6 Regions of interest

Protracted decision-making is implemented by a distributed brain network: in the context of a perceptual decision, sensory cortex encodes input signals ("evidence samples"), which association cortex accumulates over time into a decision variable, and motor regions translate the decision into a behavioral act [18, 20, 22, 23]. We ensured that our region of interest (ROI) selection covered this distributed brain network completely. Specifically, from the Wang et al. (2015) atlas [183], we selected all regions and grouped the individual regions into the following clusters [187]: V1, primary occipital cortex; V2-4, early occipital regions V2, V3 and V4; V3A/B, dorsal occipital cortex regions V3A and V3B; IPS0/1, intraparietal sulcus regions 0 and 1; IPS2/3, intraparietal sulcus areas 2 and 3; Vent. Occ, ventral occipital cortex regions VO1 and VO2; PHC, parahippocampal cortex regions PCH1 and PCH 2; MT+, temporal occipital cortex regions MT and MST; Lat. Occ, lateral occipital cortex regions LO1 and LO2; FEF, frontal eye fields. We selected three regions that were previously shown to exhibit choice-predictive lateralized activity [33]: IPS/PostCeS, intraparietal sulcus / postcentral sulcus; M1, hand area of primary motor cortex. Finally, based on the Glasser et al., 2016 atlas [182] we created the following clusters: dlPFC, dorsolateral prefrontal cortex (regions: "8C", "8Av", "i6-8", "s6-8", "SFL", "8BL", "9p", "9a", "8Ad", "p9-46v", "a9-46v", "46", "9-46d"); premotor, premotor cortex (regions: "55b", "6d", "6a", "FEF", "6v", "6r", "PEF"); pmPFC, posterior medial frontal cortex (regions: "SCEF", "p32pr", "a24pr", "a32pr", "p24"); vmPFC, ventral medial frontal cortex (regions: "p32", "s32", "a24", "10v", "10r", "25"); amPFC, anterior medial frontal cortex (regions: "d32", "8BM", "9m"). We excluded vertices from the Glasser at al. (2016) clusters that were already included in regions from Wang et al. (2015) atlas and those taken from de Gee et al. (2017) – for example, vertices corresponding to the location of FEF.

3.4.7 Analysis and modeling of choice behavior

RT was defined as the time from stimulus offset until the button press.

Signal-detection theoretic modeling The signal detection metrics sensitivity (d') and criterion (c) [24] were computed separately for each of the pupil bins. 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.

Drift diffusion modeling We used the HDDM toolbox [77] to assess the trial-by-trial, linear relationship between pupil responses and the drift (Figure 9C, Figure 9–figure supplement 1). Specifically, in the first hierarchical drift diffusion regression model we modeled single-trial drift as the following linear combination:

$$v = \beta_0 1 + \beta_1 S + \beta_2 P \quad (1)$$

where v was the single-trial drift, S was a binary vector describing the stimulus identity (1, signal+noise; 0, noise), P was a vector of the single-trial pupil responses. The fit parameters quantified how the single-trial drift was affected by the overall drift bias (i.e., mean drift criterion across trials, regardless of brainstem response, β_0), the overall drift rate (i.e., mean stimulus-dependent drift across trials, β_1), and pupil responses β_2. The parameters starting point, boundary separation, non-decision time and drift rate variability were also included in the model but did not vary pupil responses.

In the second hierarchical drift diffusion regression model we additionally modeled the single-trial drift as a function of single-trial alpha-band activity in primary visual cortex:

$$v = \beta_0 1 + \beta_1 S + \beta_2 P + \beta_3 A \quad (2)$$

where A was a vector of the single-trial alpha-band activity in V1.

These variants of the hierarchical drift diffusion regression model were fitted to behavioral and physiological data using Markov-chain Monte Carlo sampling as implemented in the HDDM toolbox [77]. 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 [77, 160, 161]. We ran 3 separate Markov chains with 25000 samples each. Of those, 2500 were discarded as burn-in. Individual parameter estimates were then estimated from the posterior distributions across the resulting 22500 samples. All group-level chains were visually inspected to ensure convergence. Additionally, we computed the Gelman-Rubin $\hat{R}$ statistic (which compares within-chain and between-chain variance) and checked that all group-level parameters had an $\hat{R}$ between 0.99-1.01.

To check for the linearity of the modeled relationships, we also fitted a regular (non-regression) version of the drift diffusion model to RT distributions for the separate responses (termed "stimulus coding" in [77]). We allowed the following parameters to vary with pupil response bin or V1 alpha-band activity bin: (i) the starting point of accumulation process; (ii) the separation between both bounds (i.e. response caution); (iii) the mean drift rate across trials; (iv) drift bias (an evidence independent constant added to the drift); (v) the non-decision time (sum of the latencies for sensory encoding and motor execution of

the choice); drift rate variability was included in the model, but did not vary with pupil response bin or V1 alpha-band activity bin. The fitting procedure was otherwise the same as above.

Mediation analysis We used two approaches for the mediation: one hierarchical approach to model the interaction between pupil responses, alpha-band activity in V1 and drift bias (Fig. 5), and one frequentist approach to model the interaction between pupil responses, alpha-band activity in various cortical regions and choice (Fig. S6).

We will first describe the frequentist approach. We fitted the following model based on standard mediation path analysis:

$$M = \alpha_0 1 + \alpha_1 S + \alpha_2 P \tag{3}$$

$$Yes = \beta_0 1 + \beta_1 S + \beta_2 P + \beta_3 M \tag{4}$$

where *Yes* is the predicted probability that the subject made a "yes"-choice on that trial, S was a binary vector describing the stimulus identity (1, signal+noise; 0, noise), M was a vector of single-trial alpha-band activity in a given cortical ROI, P was a vector of single-trial pupil responses, and α_0, α_1, α_2, β_0, β_1, β_2 and β_3 are the free parameters of the fit. The parameters were fit using freely available R-software [188]. Note that equation 4 is a logistic regression. The a-path was given by α_2; the b-path was given by β_3; the c-path (median effect) was given by a-path $\times$ b-path.

In the hierarchical approach we fitted the equations 2 and 3. Equation 2 was fitted using the hierarchical drift diffusion regression model described above. Equation 3 was refitted with a hierarchical linear regression approach (we included random intercepts and slopes, as did the drift diffusion regression model) using Bambi and PyMC3 [189, 190]. The c-path (median effect) was again given by the a-path $\times$ b-path, which now involved multiplying the two posterior distributions.

3.4.8 Statistical comparisons

We used the paired-sample t-test to test for significant differences between behavioral measures for the high vs. low pupil response bin (Fig. 1). We used the Wilcoxon signed-rank test to test for regression coefficients being significantly different from 0 (Figs. 4 and S6). On the time-frequency spectra, we used cluster-based permutation testing [191] to find clusters across time and frequency that differed for our contrast of interest (Figs. 3, 5, S3 and S4). In all other analyses (Figs. 2, 5, S2 and S5) we performed statistics directly on the group-level posterior distribution, which was made up of all the MCMC samples. Such a posterior distributions was either directly compared to 0 (the p-value quantifying the fraction of samples being bigger or smaller than 0), or to another posterior distribution (the p-value quantifying the overlap between the distributions).

3.4.9 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].

3.5 Acknowledgements

We thank Christiane Reissmann, Karin Deazle and Niels Kloosterman for help with subject recruitment and data acquisition, Roger Zimmermann for technical assistance, 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).

3.6 Supplementary Figures

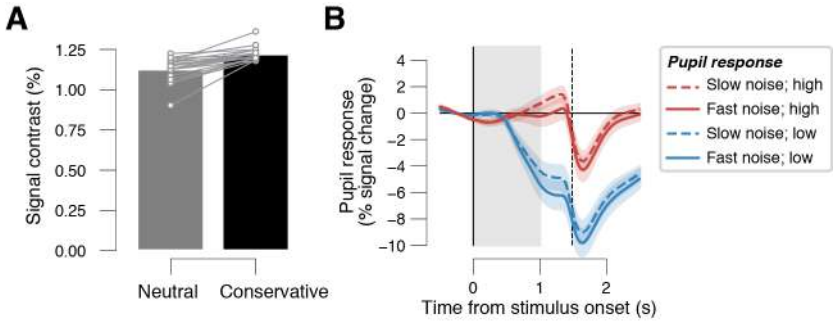


Figure S1: (A) Mean contrast values of the online staircases for the neutral and conservative condition. Data points, individual subjects. **(B)** Overall pupil response time courses in for the highest (33% of trials) and lowest (33% of trials) pupil derivative defined bins, separately for the two noise conditions. Shading, s.e.m.. All panels: Group average (N = 20).

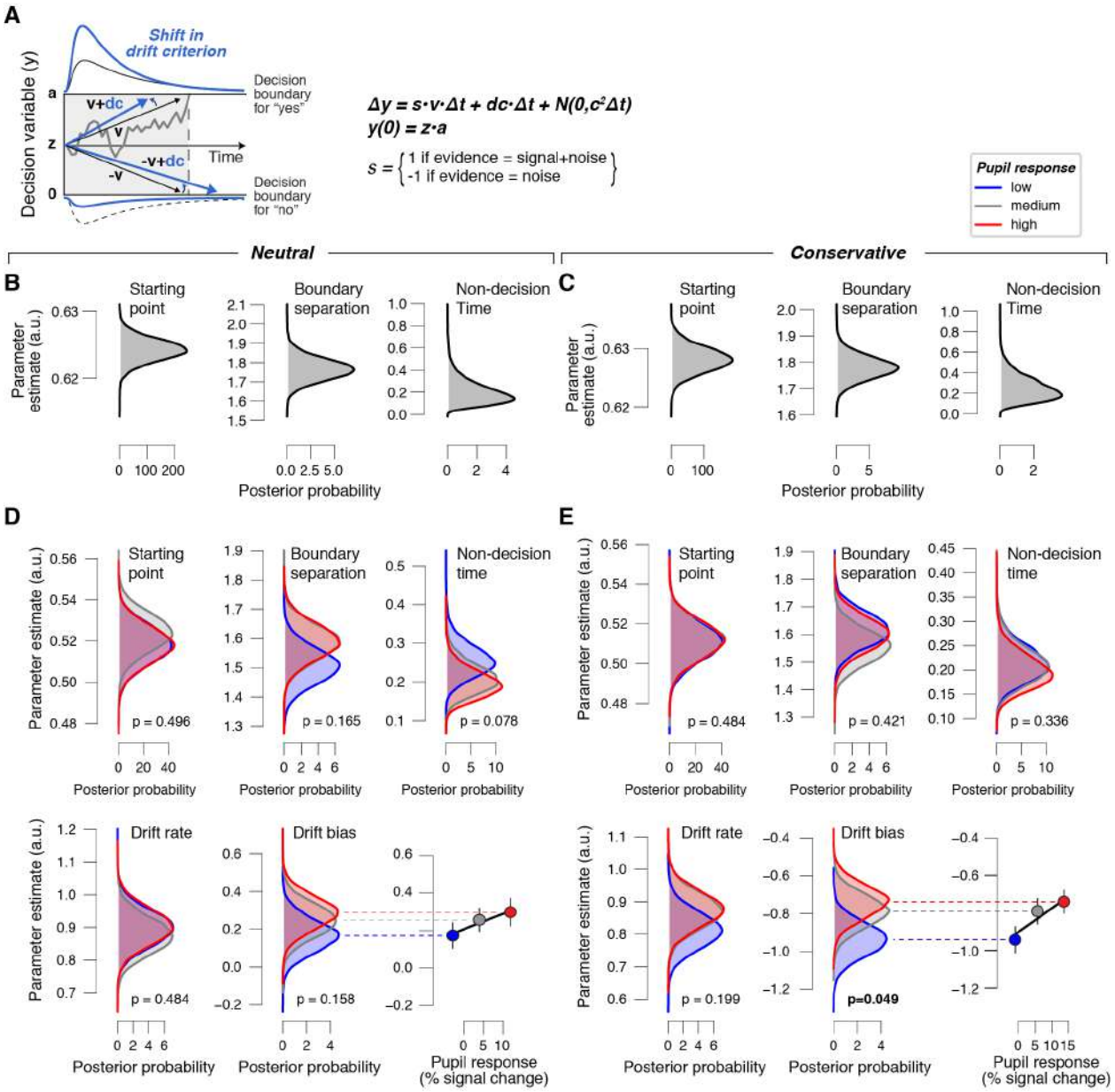


Figure S2: **(A)** Schematic of drift diffusion model accounting for choices, and their associated RTs. Solid lines, RT distributions. In the equation, v is the drift rate. **(B,C)** Group-level posterior probability densities of the remaining parameters of the hierarchical drift diffusion regression model (see Fig. 2; Materials and Methods). The parameters starting point, boundary separation, non-decision time and drift rate variability were included in the model, but varied only with condition, and not with phasic arousal. **(D,E)** Group-level posterior probability densities of all drift diffusion model parameters, when each estimated separately for three pupil-defined bins (Materials and Methods). P-values indicate the overlap between the posterior probability densities of highest (red) and lowest (blue) pupil bin. Correlation plots: relationship between drift bias estimates and pupil response bin; error bars, s.e.m. All panels: Group average (N = 20).

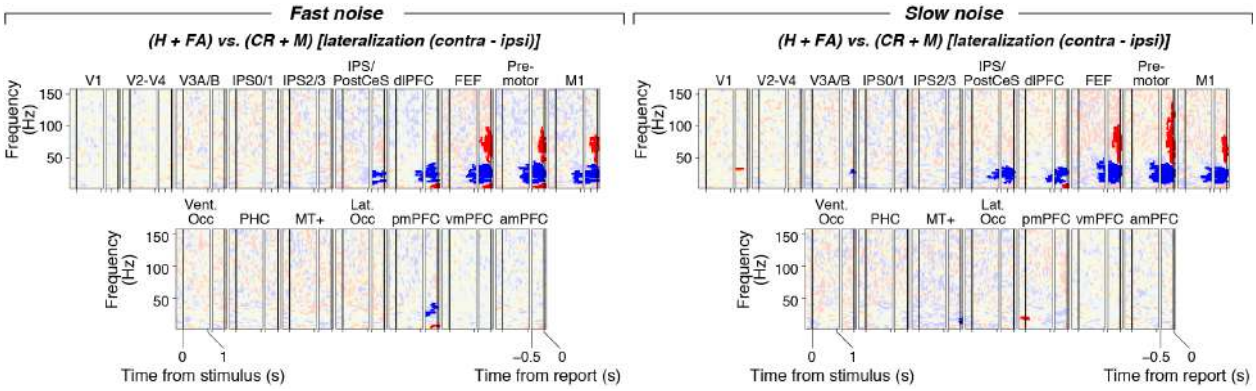


Figure S3: As Fig. 3B, but after computing the choice contrast as follows: (hit + false alarm) – (correct reject + miss), effectively factoring out effect of stimulus identity and accuracy.

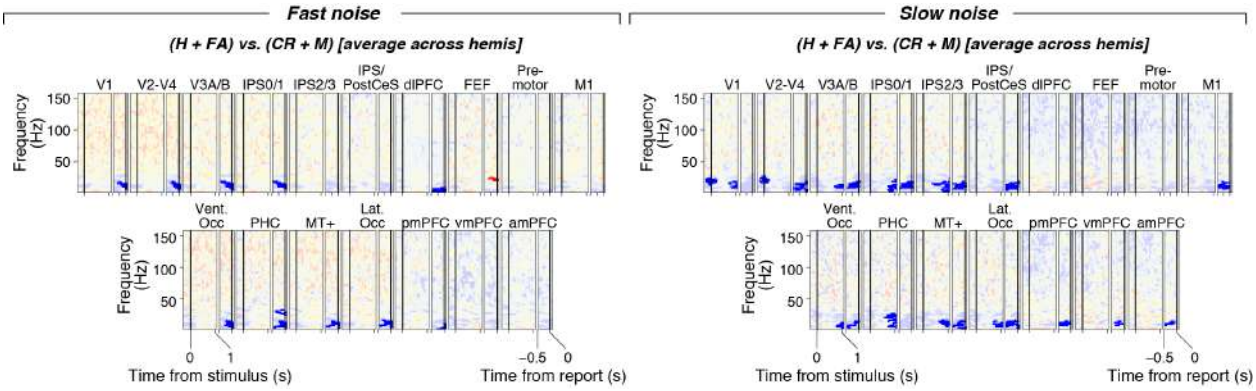


Figure S4: As Fig. 4A, but after computing the choice contrast as follows: (hit + false alarm) – (correct reject + miss), effectively factoring out effect of stimulus identity and accuracy.

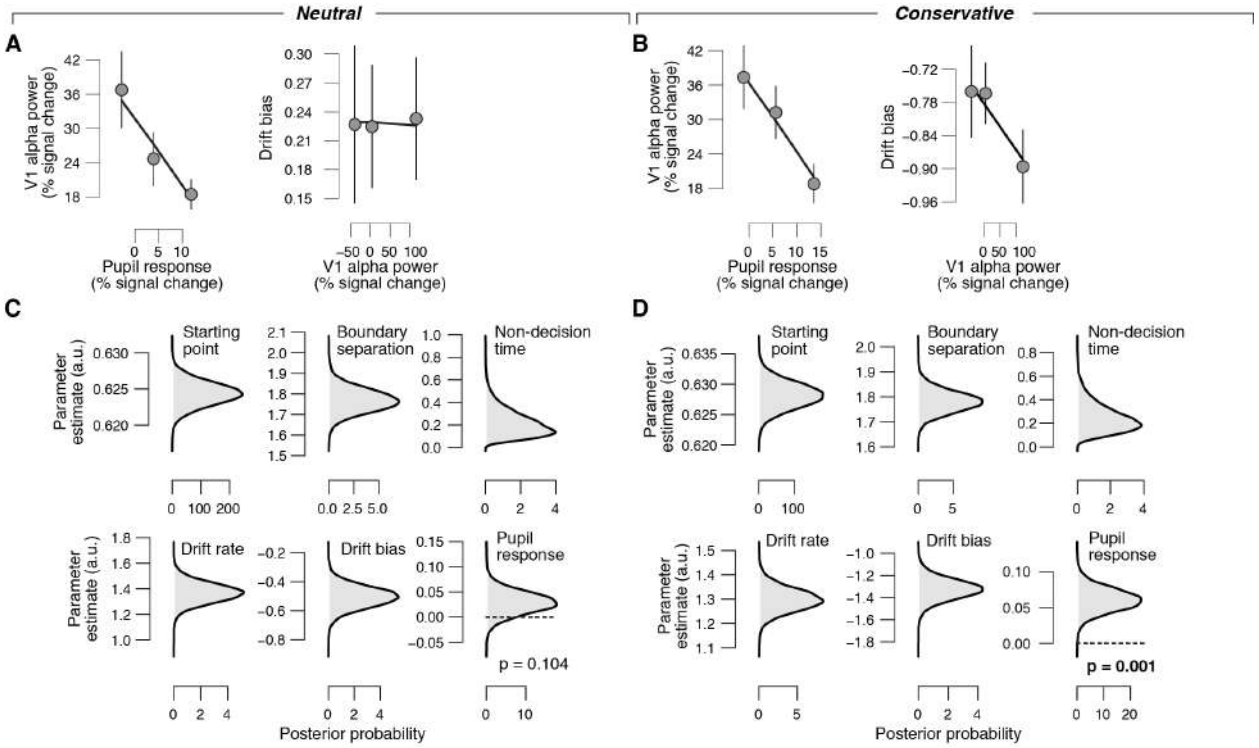


Figure S5: **(A)** Left: relationship between alpha-band activity in V1 and pupil responses in the neutral condition. Right: relationship between drift bias and alpha-band activity in V1 in the neutral condition. Error bars, s.e.m. **(B)** As panel A, but for the conservative condition. **(C)** Group-level posterior probability densities in the neutral condition for means of parameters in the hierarchical drift diffusion regression model. We used the model to assess the trial-by-trial, linear relationship between phasic arousal and alpha-band activity in V1 and the single-trial drift (Materials and Methods). **(D)** As panel C, but for the conservative condition. All panels: Group average (N = 20).

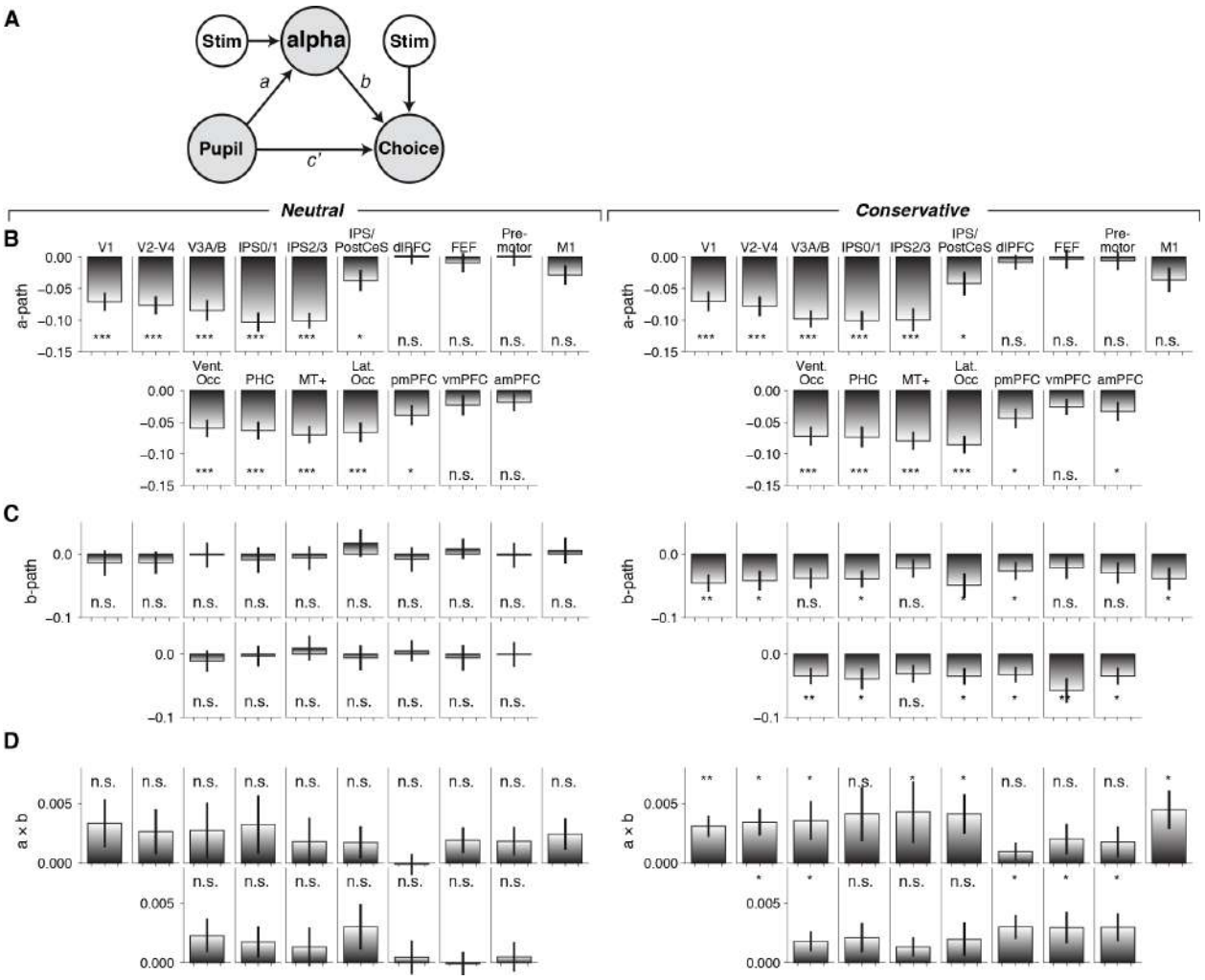


Figure S6: **(A)** Schematic of mediation analysis (see Materials and Methods). Pupil, vector of trial-wise pupil responses; alpha, vector of trial-wise alpha-band activity measures in a given ROI (Materials and Methods); stim, stimulus (signal+noise vs. noise); choice, response ("yes" vs. "no"). All arrows are regressions. **(B)** Coefficients of linear relationship between trial-wise pupil responses and alpha-band activity (a-path) for the selected ROIs. Stimulus identity (0, noise; 1, signal+noise) was added as a covariate (Materials and Methods). **(C)** Coefficients of the logistic regression of choice (0, "no": 1, "yes") on trial-wise pupil responses and alpha-band (b-path) in the selected ROIs. Stimulus identity was added as a covariate. **(D)** Coefficients of the indirect path a $\times$ b (Materials and Methods). All panels: Group average (N = 20); error bars, s.e.m.; stats, Wilcoxon signed-rank test.

Status: published
Colizoli O, de Gee JW, Urai AE & Donner TH (2018). *Scientific Reports,* 8:13702.

4

Task-evoked pupil responses reflect internal belief states

Conceptualization by OC, JWG, AEU, THD; Investigation by OC, JWG; Formal analysis by OC, JWG, AEU; Resources by JWG, THD; Writing–original draft by OC, JWG, AEU, THD; Writing–review and editing by OC, JWG, AEU, THD.

Perceptual decisions about the state of the environment are often made in the face of uncertain evidence. Internal uncertainty signals are considered important regulators of learning and decision-making. A growing body of work has implicated the brain's arousal systems in uncertainty signaling. Here, we found that two specific computational variables, postulated by recent theoretical work, evoke boosts of arousal at different times during a perceptual decision: decision confidence (the observer's internally estimated probability that a choice was correct given the evidence) before feedback, and prediction errors (deviations from expected reward) after feedback. We monitored pupil diameter, a peripheral marker of central arousal state, while subjects performed a challenging perceptual choice task with a delayed monetary reward. We quantified evoked pupil responses during decision formation and after reward-linked feedback. During both intervals, decision difficulty and accuracy had interacting effects on pupil responses. Pupil responses negatively scaled with decision confidence prior to feedback and scaled with uncertainty-dependent prediction errors after feedback. This pattern of pupil responses during both intervals was in line with a model using the observer's graded belief about choice accuracy to anticipate rewards and compute prediction errors. We conclude that pupil-

4.1 Introduction

Many decisions are made in the face of uncertainty about the state of the environment. A body of evidence indicates that decision-makers use internal uncertainty signals for adjusting choice behavior [125, 156, 192], and deviations between expected and experienced rewards for learning [193, 194]. The brain might utilize its arousal systems to broadcast such computational variables to circuits implementing inference and action selection [6, 193, 195–197].

Recent theoretical work postulates two variables at different moments during a challenging perceptual decision [127, 156, 158]: (i) decision confidence before feedback (i.e., the internally estimated probability of a choice being correct, given the available evidence) and (ii) prediction error (i.e., the difference between expected and experienced reward) after receiving feedback. Critically, and different from previous work on reinforcement learning [12, 195, 196], the prediction error signals depend on graded internal confidence [127] rather than on the categorical stimulus identity (see Model Predictions below). Such internal belief states have been incorporated in more recent models of reinforcement learning [198]. Building on previous work implicating arousal in uncertainty monitoring [74, 125, 199, 200], we here asked whether these two computational variables would evoke responses of central arousal systems.

It has long been known that the pupil dilates systematically during the performance of cognitive tasks, a phenomenon referred to as task-evoked pupil response [71, 201–207]. Physiological work indicates that non-luminance mediated changes in pupil diameter are closely linked to central arousal state [33, 34, 39, 76]. We quantified pupil responses during a perceptual choice task combined with reward-linked feedback, analogous to the task used in recent monkey work on uncertainty and prediction errors10. We then compared pupil responses before and after reward-linked feedback to predictions derived from alternative computational models of the internal variables encoded in the brain. Our goal was to (i) replicate the previously found scaling of pupil responses with decision uncertainty before feedback3 and (ii) test for the same scaling of pupil responses after feedback, as observed for dopamine neurons [127].

4.2 Results

We monitored pupil diameter in 15 human participants performing an up vs. down random dot motion discrimination task, followed by delayed reward-linked feedback (Fig. 1). The random dot motion task has been widely used in the neurophysiology of perceptual decision-making [2, 22]. Importantly, our version of the task entailed long and variable delays between decision formation and feedback, enabling us to obtain independent estimates of the pupil responses evoked by both of these events. We titrated the difficulty of the decision (by varying the evidence strength, or motion coherence, see Methods), so that observers performed at 70% correct in 2/3 of the trials in one condition ('Hard') and at 85% correct in 1/3 of the trials in the other condition ('Easy'). Correct vs. error feedback was presented after choice

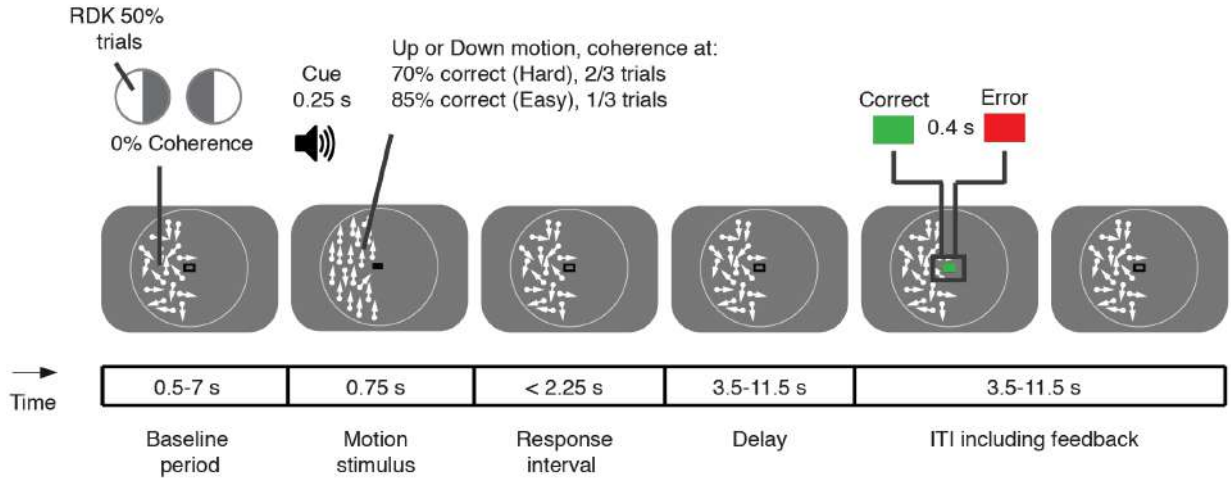


Figure 1: Perceptual choice task with delayed reward. Random dot kinematograms (RDK) were presented in one half of the visual field during each block of trials (counterbalanced). Random motion (0% coherence) was presented throughout all intervals except for the 'motion stimulus' interval, during which the RDKs to be discriminated were shown, prompted by an auditory cue (250 ms). Motion coherence of the stimulus varied from trial to trial, yielding a Hard and an Easy condition. A change back from a closed to an open rectangle in the fixation region (constant luminance) prompted subjects' choice ('Response interval'). After a variable delay (3.5–11.5 s) following the choice, feedback was presented that was coupled to a monetary reward (see main text). The white circle surrounding the RDKs is for illustration only and was not present during the experiment.

and converted into a monetary reward, based on the average performance level across a block (25 trials), as follows: 100% correct yielded 10 Euros, 75% yielded 5 Euros, chance level (50% correct) yielded 0 Euros. The total reward earned was presented on the screen to participants at the end of each block.

4.2.1 Model predictions

We used two computational models based on signal detection theory [24] to generate qualitative predictions for the behavior of internal signals before and after reward feedback that might drive pupil-linked arousal (Fig. 2a, see Materials and Methods for details). Both models assumed that observers categorize the motion direction based on a noisy decision variable, which in turn depended on the stimulus strength (motion coherence), the stimulus identity (Up or Down), and on internal noise. The models' choices were governed by comparing this noisy decision variable to zero, ensuring no bias towards one over the other choice.

The two models differed in how confidence was defined. Here, with confidence we refer to the observer's internally estimated probability that a choice was correct given the available evidence [158]. Because choice accuracy was coupled to a fixed monetary reward in our experiment (see above), confidence equaled an ideal observer's internally estimated probability of obtaining the reward, in other words, reward expectation. In the 'Belief State Model', confidence was computed as the absolute distance between the decision variable (depending on the stimulus identity, stimulus strength, and internal noise) and the

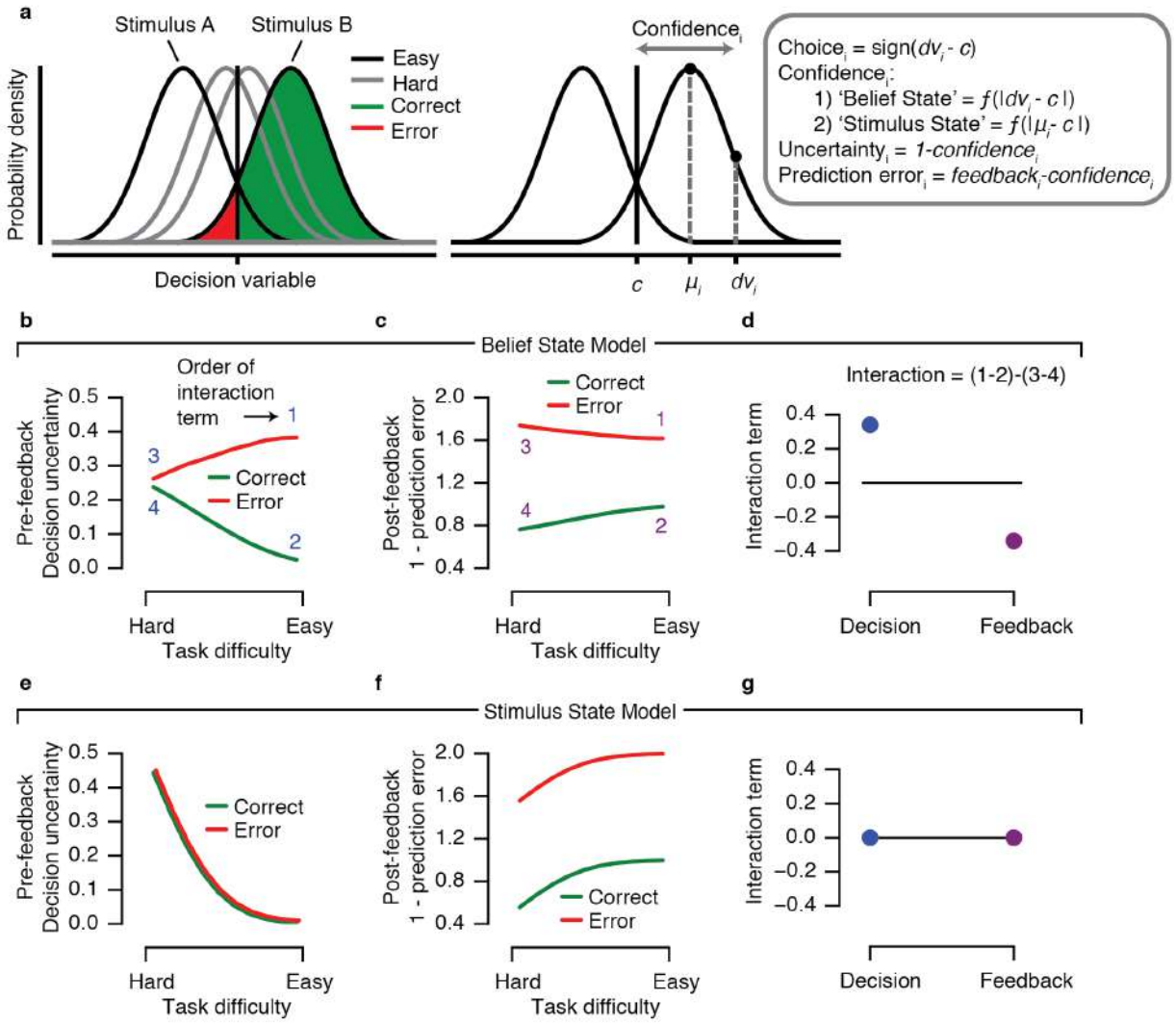


Figure 2: Alternative predictions for internal signals during pre- and post-feedback intervals of the task. **(a)** Computations underlying choice, confidence, uncertainty and prediction error. Repeated presentations of a generative stimulus produce a normal distribution of internal decision variables (dv) due to the presence of internal noise, which is centered around the generative stimulus (μ). In this model, confidence is defined as the single-trial distance between dv and c, the internal decision bound. Prediction errors are computed by comparing experienced reward (i.e. feedback) with the observers' expected outcome. **(b-d)** Computational variables were simulated for every trial, then averaged separately for Correct and Error conditions for each level of task difficulty (in this case, motion coherence). Belief State Model predictions. **(e-g)** Stimulus State Model predictions. **(b,e)** Decision uncertainty (complement of confidence) as function of task difficulty during pre-feedback interval. **(c,f)** Prediction error as function of task difficulty during post-feedback interval. **(d,g)** Interaction term computed as (Easy Error - Easy Correct) - (Hard Error - Hard Correct). See main text for model details.

decision criterion (i.e., zero) (Fig. 2a; see Methods and [156]). By contrast, in the 'Stimulus State Model', confidence was computed as the absolute distance between the physical stimulus value (i.e. physical stimulus identity times stimulus strength) and the criterion (zero). In both models, reward prediction error was computed as the difference between the confidence and the reward-linked feedback. Thus, in the Belief State Model, the observer's internal belief about the state of the outside world (encoded in the noisy

decision variable) determined both reward expectation (i.e., confidence) and reward prediction error; in the Stimulus State Model, these computational variables did not depend on the observer's internal belief, but only on the strength and identity of the external stimulus.

We simulated these two models to derive qualitative predictions that distinguished between their internal signals. To this end, we computed confidence and reward prediction errors at the level of individual trials (see above) and then collapsed these single-trial signals within each Accuracy and Difficulty condition. The rationale was that the interaction between conditions (defined as [Easy Error - Easy Correct] - [Hard Error - Hard Correct]) most clearly dissociated between the predictions generated from both models (Fig. 2b–g).

Previous pupillometry work on a similar task showed that pre-feedback pupil responses scaled with decision uncertainty (i.e. the complement of decision confidence) [125]. We thus generated predictions for decision uncertainty during the pre-feedback interval (Fig. 2b,e) and, by analogy, for the complement of the prediction error during the post-feedback interval (Fig. 2c,f).

The critical observation is that the Belief State Model predicts a positive Accuracy × Difficulty interaction pre-feedback, and a negative interaction post-feedback (Fig. 2d). This pattern is consistent with predictions from a reinforcement learning model based on a partially observable Markov decision process (POMDP) [127]. In contrast, the Stimulus State Model does not predict an Accuracy × Difficulty interaction either pre- or post-feedback (Fig. 2g). This pattern is consistent with traditional reinforcement learning models [12, 195, 196].

Previous work on perceptual choice has shown that reaction time (RT) scales with decision uncertainty [125, 208, 209], in line with the Belief State Model. The same was evident in the present data: there was a main effect of accuracy, $F(1,14) = 51.57$, $p < 0.001$, and difficulty, $F(1,14) = 19.53$, $p < 0.001$, as well as an interaction effect of both, $F(1,14) = 34.95$, $p < 0.001$, on RT (see Supplementary Fig. S1, compare with Fig. 2b), in line with the Belief State Model. This indicates that, in our current data, a graded, noisy decision variable similar to the one postulated by the Belief State Model was encoded and used for the decision process. We next tested which of the two models better reflected responses of pupil-linked arousal systems. We analyzed pupil responses as a function of motion strength and choice correctness for the two critical intervals of the trial: the phase of reward anticipation before feedback, as in previous work [125], and critically, the phase of reward prediction error signaling after feedback.

4.2.2 Sustained pupil response modulations during pre- and post-feedback intervals

The pupil responded in a sustained fashion during both intervals: after the onset of the motion stimulus and locked to the observers' reported choice (i.e., pre-feedback) and post-feedback (Fig. 3a, blue and purple lines). The pupil response remained elevated during feedback anticipation, long after stimulus processing (maximum of 3 s, 0.75 s stimulus duration plus response deadline of 2.25 s, see Fig. 1). Upon feedback presentation, the pupil initially constricted due to the presentation of the visual feedback stimulus (see Supplementary Fig. S2) and then dilated again to a sustained level for the remainder of the post-

feedback interval. Please note that we subtracted the pupil diameter during the pre-feedback period from the feedback-locked responses (see Methods), so as to specifically quantify the feedback-evoked response.

For comparison, we measured, in the same participants (separate experimental blocks), pupil responses evoked during a simple auditory detection task (button press to salient tone), which did not entail prolonged decision processing and feedback anticipation (see Methods). The resulting response, termed 'impulse response function' (IRF) for simplicity, was more transient than those measured during the main experiment: the IRF returned back to the pre-stimulus baseline level after 3 s (Fig. 3a, compare grey IRF with the blue line). Thus, the sustained elevations of pupil diameter observed beyond that time in the main experiment reflected top-down, cognitive modulations in pupil-linked arousal due to decision processing and reward anticipation (for the responses locked to the onset of the choice), or due to reward processing (for the feedback-locked responses). To quantify the amplitude of these cognitive modulations of the pupil response, we collapsed the pupil response across the time window 3–6 s from the choice (for pre-feedback interval) or from the feedback (post-feedback interval; see gray shaded area in Fig. 3a). For the cognitive modulations during the pre-feedback interval, we further extracted the mean pupil response values in the 500 ms before the feedback (gray shaded area in Fig. 3b).

4.2.3 Interacting effects of decision difficulty and accuracy on evoked pupil responses

The sustained pupil responses during both the intervals, pre- and post-feedback, scaled in line with the predictions from the Belief State Model, not the Stimulus State Model (compare Fig. 3d–f with Fig. 2b–d). First, pupil responses during both intervals were overall larger on Error than Correct trials (Fig. 3b,c). The Stimulus State Model did not predict any difference between the two categories during the pre-feedback interval, because this model was only informed by external information (motion stimulus or feedback), not by noisy internal states. The larger pupil responses during errors in the pre-feedback interval were in line with previous results3, supporting the idea that arousal state between choice and feedback reflects the observer's decision uncertainty.

Second, the sustained pupil responses during both intervals exhibited a pattern of interactions between decision difficulty and accuracy as predicted by the Belief State Model but not the Stimulus State Model (compare Fig. 3d to Fig. 2d and Fig. 2g). Hereby, the interaction was defined as (Easy Error - Easy Correct) - (Hard Error - Hard Correct). Specifically, the Belief State Model predicted a significant interaction of opposite sign for both intervals (Fig. 2d, compare blue and purple dots). That same pattern was evident in the time course of the interaction term in the pupil response. During both intervals, the interaction terms were significant, with opposite signs: positive during the pre-feedback interval and negative during the post-feedback interval (Fig. 3d, blue and purple bars). Consequently, the interaction terms were significantly different from one another throughout the entire part of the sustained pupil response (Fig. 3d, black bar).

Finally, also the full pattern of sustained pupil responses for the Hard vs. Easy and Correct vs. Error conditions in both trial intervals (Fig. 3e,f) resembled the pattern predicted by the Belief State Model

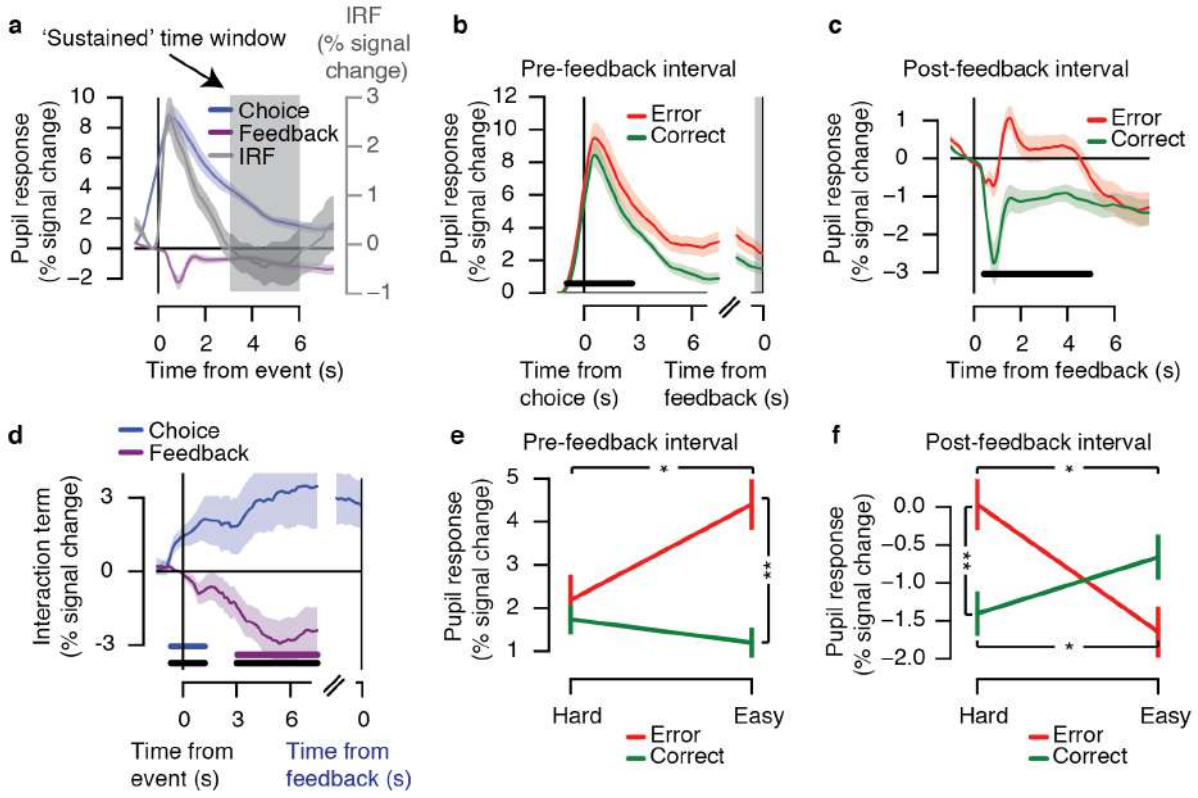


Figure 3: Pupil responses before and after feedback reflect observers' belief state. (a) Pupil responses locked to the observer's reported choice (blue) and locked to feedback (purple). Shown for comparison is the pupil 'impulse response' from same participants (IRF, see main text). Grey shading indicates sustained time window (3-6 s), in which IRF returned to baseline. **(b)** Evoked pupil responses for Correct and Error trials in pre-feedback interval. Black bar indicates correct vs. error effect, $p < 0.05$ (cluster-based permutation test). **(c)** Evoked pupil responses for Correct and Error trials in post-feedback interval. Black bar indicates correct vs. error effect, $p < 0.05$ (cluster-based permutation test). **(d)** Interaction term (Easy Error - Easy Correct) - (Hard Error - Hard Correct) for choice-locked (blue, coinciding with observers' reported choice) and feedback-locked (purple) responses. Horizontal bars indicate effect, $p < 0.05$ (cluster-based permutation test): blue bar indicates choice-locked response tested against 0; purple bar indicates feedback-locked response tested against 0; black bar indicates difference in interaction between both responses. **(e)** Mean response in the pre-feedback interval (500 ms preceding feedback), as a function of difficulty and accuracy. **(f)** Mean response (in sustained time window) during the post-feedback interval, as function of difficulty and accuracy. Error bars, standard error of the mean (N = 15). *p < 0.05, **p < 0.01, ***p < 0.001.

(Fig. 2b,c). In the sustained window during the post-feedback interval, there was a significant interaction between difficulty and accuracy (Fig. 3f, $F(1,14) = 9.31$, $p = 0.009$; post hoc comparisons: Hard Error vs. Hard Correct, $p = 0.001$; Easy Error vs. Easy Correct, $p = 0.174$; Hard Error vs. Easy Error, $p = 0.037$; Hard Correct vs. Easy Correct, $p = 0.031$). The sustained window before feedback exhibited a trend towards an interaction ($F(1,14) = 4.12$, $p = 0.062$). This effect became stronger during the pre-feedback interval: in the 500 ms window just before (and locked to) the feedback delivery (grey window in Fig. 3b), the interaction was significant (Fig. 3e; $F(1,14) = 6.66$, $p = 0.022$; post hoc comparisons: Hard Error vs. Hard Correct, $p = 0.500$; Easy Error vs. Easy Correct, $p = 0.006$; Hard Error vs. Easy Error, $p = 0.037$; Hard Correct vs. Easy Correct, $p = 0.060$). For all subsequent analyses, we focus on this interval 500 ms before feedback to probe into participants' reward anticipation, referring to this time window as the 'pre-feedback interval'.

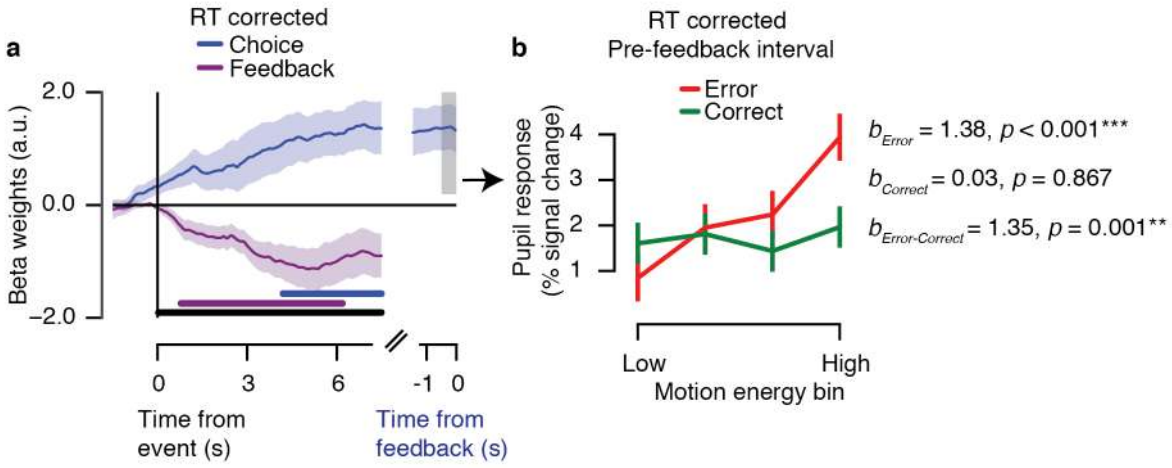


Figure 4: Pupil responses before feedback reflect observers' belief state even when controlling for RT and motion energy fluctuations. **(a)** Time course of belief state scaling in the pupil, computed as trial-by-trial regression of RT-corrected pupil dilation onto motion energy strength. The interaction term (beta weights Error - beta weights Correct) is shown for choice-locked (blue, coinciding with onset of the choice) and feedback-locked (purple) responses. Horizontal bars indicate effect, p < 0.05 (cluster-based permutation test): blue bar indicates choice-locked response tested against 0; purple bar indicates feedback-locked response tested against 0; black bar indicates difference in interaction between both responses. **(b)** Mean response in pre-feedback time window (-0.5–0 s) as a function of difficulty and accuracy. Absolute motion energy was divided into four equally sized bins (per participant) for visualization. Error bars, standard error of the mean (N = 15). *p < 0.05, **p < 0.01, ***p < 0.001.

In sum, in this perceptual choice task, sustained pupil responses during both reward anticipation (pre-feedback) as well as after reward experience (post-feedback) were qualitatively in line with the predictions from a model of reward expectation and prediction error, in which the computation of these internal variables depended on internal belief states. The results from all main figures are only based on trials with long delay intervals (≥7.5 s) between choice and feedback, and between feedback and the subsequent trial, in order to minimize possible contamination of evoked pupil responses by the next event (i.e., feedback or the next trial's cue; see Methods). We found the same pattern of results when performing the analyses on all trials (Supplementary Fig. S3).

4.2.4 Control analysis for confounding effects of variations of RT and motion energy

In the current study, as in previous work using a similar perceptual choice task [125], both RT and pre-feedback pupil dilation scaled with the decision uncertainty signal postulated by the Belief State Model. Indeed, RTs were significantly correlated to pre-feedback pupil responses in the pre-feedback window (-0.5–0 s) across all trials, $r(13) = 0.12$, $p < 0.001$, and within the following conditions: Hard Error, $r = 0.11$, $p = 0.001$; Hard Correct, $r = 0.09$, $p < 0.001$; Easy Correct, $r = 0.16$, $p < 0.001$, but not within the Easy Error condition, $r = 0.07$, $p = 0.223$.

While this association was expected under the assumption that RT and pupil dilation were driven by internal uncertainty signals [125], the association also raised a possible confound. Arousal drives pupil

dilation in a sustained manner throughout decision formation [33, 40, 42]. The peripheral pupil apparatus for pupil dilation (nerves and smooth muscles) has temporal low-pass characteristics. Consequently, trial-to-trial variations in decision time (the main source of RT variability) can cause trivial trial-to-trial variations in pupil dilation amplitudes, simply due to temporal accumulation of a sustained central input of constant amplitude but variable duration [33, 40]. Then, pre-feedback pupil response amplitudes may have reflected RT-linked uncertainty, but without a corresponding scaling in the amplitudes of the neural input from central arousal systems. Note that this concern applied only to the pre-feedback pupil dilations, not the post-feedback dilations, which were normalized using the pre-feedback interval as a baseline (see above). Another concern was that trial-by-trial fluctuations in motion energy, caused by the stochastically generated stimuli (see Methods) contributed to behavioral variability within the nominally Easy and Hard conditions.

Our results were not explained by either of those confounds (Fig. 4). To control for both of them conjointly, we removed the influence of trial-to-trial variations in RT (via linear regression) from the pre-feedback pupil responses. And we used motion energy filtering [125, 210] to estimate each trial's sensory evidence strength. We finally regressed the RT-corrected pupil time courses onto evidence strength (absolute motion energy), separately for the Error and Correct trials. The interaction term was defined as the difference in beta weights for the Error vs. Correct trial regressions. In this control analysis, the critical interaction effect was significant during both the pre-feedback and post-feedback time courses ($ps < 0.05$, cluster-based permutation test; Fig. 4a). The interaction terms furthermore differed between intervals ($p < 0.05$, cluster-based permutation test; Fig. 4a). When regressing mean RT-corrected pupil responses in the pre-feedback time window onto evidence strength, the critical interaction term (i.e. beta weights) within the pre-feedback window still reflected decision uncertainty (Fig. 4b; $M = 1.35$, $STD = 1.81$, $p = 0.001$). In sum, while trial-to-trial variations in RT and motion energy explained some variance in the pupil responses, the key patterns of the pupil responses diagnostic of modulation by belief states were robust even when controlling for these parameters.

4.2.5 Relationship to Urai et al., 2017

Our current results from the pre-feedback interval replicate our earlier finding [125] that pupil responses in perceptual choice scale with decision uncertainty as postulated by the Belief State Model. This previous study focused on the pre-feedback responses and did not specifically assess the feedback-locked pupil responses (pupil measures were corrected with the same pre-trial baseline for the entire trial [125]). We here re-analyzed the post-feedback responses in the data from [125] for comparison (see Supplementary Fig. S4). As in our current data, post-feedback responses were larger after incorrect than correct feedback (Supplementary Fig. S4a). However, the uncertainty-dependent scaling of post-feedback responses differed: rather than a negative interaction effect (Fig. 3d), the interaction effect after feedback was positive (Supplementary Fig. S5b,c). One possible explanation for this difference may be the effect of reward-linked feedback: while participants in the current study were paid a compensation depending on their

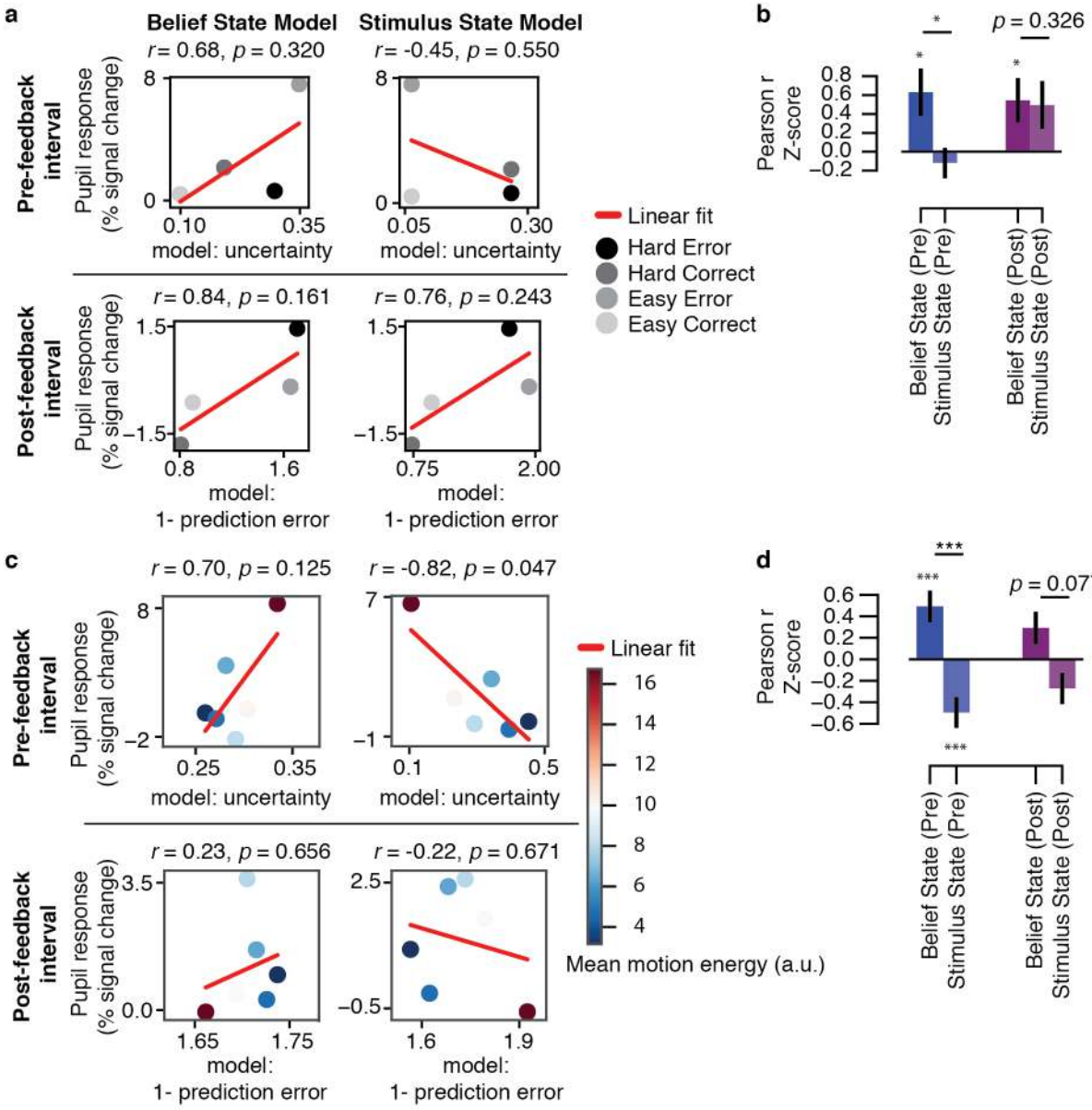


Figure 5: Model fits to pupil responses. **(a)** An example of the correlations (r) for a single subject. The four conditions of interest were defined by the Accuracy x Difficulty interaction. Easy and Hard conditions for the model parameters were averaged based on the coherence levels presented to each subject. **(b)** Group-level correlation coefficients (r) for the comparison of the model parameters and pupil responses, for the pre-feedback (Pre; -0.5-0 s) and post-feedback (Post; 3-6 s) intervals. **(c)** An example of the correlations for a single subject using model parameters simulated with motion energy (error trials only). Pupil responses were averaged within equal-sized bins based on the model parameter for each interval (6 bins). Evidence strength is represented by mean motion energy within each bin (color bar). **(d)** Group-level correlation coefficients (r) for the comparison of the model parameters (using motion energy) and pupil responses, for the pre-feedback (Pre; -0.5-0 s) and post-feedback (Post; 3-6 s) intervals (error trials only). Error bars, standard error of the mean (N = 15). **p < 0.01, ***p < 0.001.

performance, feedback in the study by [125] did not affect a monetary reward. It is thus possible that the prospect of receiving a performance-dependent monetary reward is required for the recruitment of pupil-

linked arousal systems by uncertainty-dependent prediction errors. A number of further differences between these two studies complicates a direct comparison: the behavioral task (i.e. the comparison of two intervals of motion strength vs. coarse motion direction discrimination), the short vs. long delay periods between events, and the two cohorts of participants. Despite these limitations, the difference in results between studies is potentially relevant and should be tested directly in follow-up work that eliminates the confounding factors listed above.

4.2.6 Belief State Model predicts pupil responses quantitatively better than Stimulus State Model

The data presented thus far show that the pattern of pupil responses was qualitatively in line with the Belief State Model but not with the Stimulus State Model. To this end, we used predictions from model simulations based on the group data. However, individuals differ widely in terms of the internal noise, which dissociates between the models. We next tested whether the Belief State Model provides a quantitatively superior match to the measured pupil data than the Stimulus State Model when individual estimates of internal noise are used to generate model predictions. To this end, we simulated both models using individual estimates of internal noise (Supplementary Fig. S5a and Methods). This yielded model predictions for each individual for the Accuracy $\times$ Difficulty conditions, which were qualitatively in line with predictions based on the group, but with effects that varied in their magnitude between individuals depending on their estimated internal noise (Supplementary Fig. S5b).

We predicted that those individual patterns predicted by the Belief State Model should be more similar to the measured individual pupil responses than the individual patterns predicted by the Stimulus State Model. We tested this prediction by correlating predictions of both models with the corresponding pupil responses, separately for each individual. An example for a single subject is shown in Fig. 5a, for both trial intervals. For both intervals, group-level correlations (Fig. 5b) were significant (i.e. pupil responses similar) for the Belief State Model ($ps < 0.05$), but not the Stimulus State Model ($ps > 0.05$). Further, for the pre-feedback interval, the Belief State Model correlations were significantly stronger than the Stimulus State Model ($p < 0.05$). For the post-feedback interval, there was a trend towards a stronger correlation for the Belief State Model than the Stimulus State Model ($p = 0.074$).

To perform a more fine-grained evaluation of the correspondence between model-predicted patterns and pupil responses, we used the motion energy information extracted from each trial (see previous section and Methods) rather than the categorical difficulty conditions (Easy, Hard) to generate individual model predictions. Because errors, not correct trials, qualitatively dissociate the predictions from the Belief State and Stimulus State Models (compare Fig. 2b,c with Fig. 2e,f), we restricted this control analysis to error trials (Fig. 5c,d). Again, predictions of both models were correlated to the corresponding pupil responses (6 bins of model parameters), separately for each individual. An example for a single subject is shown in Fig. 5c.

For both intervals, correlations were positive (i.e. pupil responses similar) for the Belief State model

predictions and negative (i.e. pupil responses dissimilar) for the Stimulus State model. Critically, the Belief State Model correlations were significantly larger than the Stimulus State Model in the pre-feedback interval ($p < 0.001$), again with a similar trend for the post-feedback interval ($p = 0.077$). The same held for a single-trial version of this correlation analysis, again focusing on error trials only (difference in correlation between models: $p < 0.001$ for pre-feedback; $p = 0.080$ for post-feedback).

4.3 Discussion

It has long been known that the pupil dilates systematically during the performance of cognitive tasks [71, 201–207]. The current study shows that task-evoked pupil dilation during a perceptual choice task indicates, at different phases of the trial, decision uncertainty and reward prediction error. Comparisons with qualitative model predictions showed that pupil responses during feedback anticipation and after reward feedback were modulated by decision-makers' (noise-corrupted) internal belief states that also governed their choices. This insight is consistent with a reinforcement learning model (POMDP) that incorporates graded belief states in the computation of the prediction error signals [127, 198]. In sum, the brain's arousal system is systematically recruited in line with high-level computational variables.

A number of previous studies have related non-luminance mediated pupil responses to decision-making, uncertainty, and performance monitoring [40, 73, 74, 199, 200, 211–213], but our current results move beyond their findings in important ways. First, with the exception of [125], previous studies linking uncertainty to pupil dynamics have used tasks in which uncertainty originated from the observer's environment [74, 199, 200, 211]. By contrast, in our task, decision uncertainty largely depended on the observers' internal noise, which dissociated the two alternative models of the computational variables under study (decision uncertainty and reward prediction error, Fig. 2). Second, our work went beyond the results from [125] in showing that post-feedback pupil dilation reflects belief-modulated prediction error signals during perceptual decision-making in the context of a monetary reward.

Previous work on central arousal systems and pupil-linked arousal dynamics has commonly used the dichotomy of (i) slow variations in baseline arousal state and (ii) rapid (so-called 'phasic') evoked responses [6, 11, 40, 101]. Our current results indicate that this dichotomy is oversimplified, by only referring to the extreme points on a natural continuum of arousal dynamics during active behavior. Our results show that uncertainty around the time of decision formation as well as the subsequent reward experience both boost pupil-linked arousal levels in a sustained fashion: pupils remained dilated for much longer than what would be expected from an arousal transient (Fig. 3, compare all time courses with the IRF). Even in our comparably slow experimental design, these sustained dilations lasted until long after the next experimental event. This implies that the sustained evoked arousal component we characterized here contributes significantly to trial-to-trial variations in baseline pupil diameter, which have commonly been treated as 'spontaneous' fluctuations.

Our insights are in line with theoretical accounts of the function of neuromodulatory brainstem systems implicated in the regulation of arousal [6, 197]. Recent measurements in rodents, monkeys, and

humans have shown that rapid pupil dilations reflect responses of neuromodulatory nuclei [33, 35, 39]. Neuromodulatory systems are interesting candidates for broadcasting uncertainty signals in the brain because of their potential of coordinating changes in global brain state [6, 11] and enabling synaptic plasticity in its target networks [214, 215]. While pupil responses evoked by decision tasks or micro-stimulation have commonly been associated with the noradrenergic locus coeruleus [33, 35, 36, 38, 39], these studies also found correlates in other brainstem systems [33, 35, 39]. In particular, task-evoked pupil responses during perceptual choice correlate with fMRI responses in dopaminergic nuclei, even after accounting for correlations with other brainstem nuclei (Figure 8H in [33]). Several other lines of evidence also point to an association between dopaminergic activity and non-luminance mediated pupil dilations. First, the locus coeruleus and dopaminergic midbrain nuclei are (directly and indirectly) interconnected [91, 216, 217]. Second, both receive top-down input from the same prefrontal cortical regions [91], which might endow them with information about high-level computational variables such as belief states. Third, task-evoked fMRI responses of the locus coeruleus and substantia nigra are functionally coupled (Figure 8G in [33]). Fourth, both neuromodulatory systems are implicated in reward processing [36, 216]. Fifth, rewards exhibit smaller effects on pupil dilation in individuals with Parkinson's disease than in age-matched controls, a difference that can be modulated by dopaminergic agonists [218]. Future invasive studies should establish this putative link between pupil diameter and the dopamine system.

Recordings from midbrain dopamine neurons in monkey have also uncovered dynamics on multiple timescales [102, 219], in line with our current insights into pupil-linked uncertainty signaling. Further, the pattern of pupil dilations measured in the current study matched the functional characteristics of dopamine neurons remarkably closely (specifically, the pattern of the interaction between task difficulty and accuracy in pre- and post-feedback responses) [127]. However, the pupil responses followed the complement of the computational variables (i.e., 1-confidence and 1-prediction error) and the dopamine neurons identified by [127]. It is tempting to speculate that task-evoked pupil responses track, indirectly, the sign-inverted activity of such a belief-state modulated dopaminergic system. Another alternative is that other brainstem systems driving pupil dilations [33, 35, 39] exhibit the same belief-state modulated prediction error signals as dopamine neurons.

Our current work has some limitations, but also broader implications, which might inspire future work. First, provided that participants had learned the required (constant) decision boundary, the current task did not require them to learn any environmental statistic. While a prediction error signal such as the one studied here may be essential for perceptual learning [220, 221], the importance of the pupil-linked arousal signals for learning remains speculative in the context of our experiment. Future work should address their link to learning. In particular, while decision uncertainty can also be read out from behavioral markers such as RT [125, 208, 209], no overt behavioral response is available to infer internal variables instantiated in response to feedback. Thus, our insight that the post-feedback pupil dilation reports a signal that is known to drive learning in the face of state uncertainty [198] paves the way for future studies using this autonomous marker for tracking such signals in the brain.

Another important direction for future research is the relationship between pupil-linked uncertainty signals and the sense of confidence as reported by the observer [73]. The Belief State Model we used here makes predictions about a computational variable, statistical decision confidence [158], while being agnostic about the mapping to the sense of confidence experienced or reported by the observer. Human confidence reports closely track statistical decision confidence in some experiments [209], but suffer from miscalibration in others, exhibiting over- or underconfidence [222], insensitivity to the reliability of the evidence [223], or biasing by affective value [224].

In sum, we have established that internal belief states during perceptual decision-making, as inferred from a statistical model, are reflected in task-evoked pupil responses. This peripheral marker of central arousal can be of great use to behavioral and cognitive scientists interested in the dynamics of decision-making and reward processing in the face of uncertainty.

4.4 Methods

An independent analysis of these data for the predictive power of pupil dilation locked to motor response, for perceptual sensitivity and decision criterion has been published previously25. The analyses presented in the current paper are conceptually and methodologically distinct, in that they focus on the relationship between Belief State Model predictions and pupil dilation, in particular locked to the presentation of reward feedback.

4.4.1 Participants

Fifteen healthy subjects with normal or corrected-to-normal vision participated in the study (6 women, aged 27 ± 4 years, range 23–37). The experiment was approved by the Ethical Committee of the Department of Psychology at the University of Amsterdam. All subjects gave written informed consent. All experiments were performed in accordance with the ethical guidelines and regulations. Two subjects were authors. Subjects were financially compensated with 10 Euros per hour in the behavioral lab and 15 Euros per hour for MRI scanning. In addition to this standard compensation, subjects earned money based on their task performance: 0–10 Euros linearly spaced from 50–100% accuracy per experimental session (i.e. 50% correct = 0 Euros, 75% = 5 Euros, 100% = 10 Euros). At the end of each block of trials, subjects were informed about their average performance accuracy and corresponding monetary reward. Earnings were averaged across all blocks at the end of each session.

4.4.2 Behavioral task and procedure

Subjects performed a two-alternative forced choice (2AFC) motion discrimination task while pupil dilation was measured (Fig. 1). Motion coherence varied so that observers performed at 70% correct in 2/3 of trials ('Hard') and at 85% correct in 1/3 of trials ('Easy'). After a variable delay (3.5–11.5 s) following the choice on each trial, we presented feedback that was coupled to a monetary reward (see 'Participants').

Each subject participated in one training session and four main experimental sessions (in the MRI scanner). During the training session, subjects' individual threshold coherence levels were determined using a psychometric function fit with 7 levels, 100 trials per level, 0–80% coherence. The training session took 1.5 hours and each experimental session lasted 2 hours. During the experimental sessions, stimuli were presented on a 31.55" MRI compatible LCD display with a spatial resolution of 1920 × 1080 pixels and a refresh rate of 120 Hz.

The individual coherence levels were validated at the beginning of each experimental session in practice blocks (during anatomical scans) by checking that the subject's average accuracy across a block corresponded to 75% correct. If subjects' average accuracy of a block exceeded 75%, the difficulty of the task was increased in the following block by slightly decreasing the motion coherence based on individual performance thresholds (in steps of 1% in accuracy, equally for both Hard and Easy conditions). During experimental blocks, greater motion coherence (i.e. stronger evidence strength) resulted in higher accuracy as well as faster responses. Subjects' accuracy was higher on Easy trials ($M = 88.06\%$ correct, $SD = 4.26$) compared to Hard trials ($M = 71.15\%$ correct, $SD = 3.64$), $p < 0.001$. Subjects were faster to respond on Easy trials ($M = 1.13$ s, $SD = 0.13$) compared to Hard trials ($M = 1.22$ s, $SD = 0.14$), $p < 0.001$.

Task instructions were to indicate the direction of coherent dot motion (upward or downward) with the corresponding button press and to continuously maintain fixation in a central region during each task block. Subjects were furthermore instructed to withhold responses until the offset of the coherent motion stimulus (indicated by a visual cue). The mapping between perceptual choice and button press (e.g., up/down to right/left hand button press) was reversed within subjects after the second session (out of four) and was counterbalanced between subjects. Subjects used the index fingers of both hands to respond.

Each trial consisted of five phases during which random motion (0% coherence) was presented, with the exception of the stimulus interval: (i) the pupil baseline period (0.5–7 s); (ii) the stimulus interval consisting of random and coherent motion for a fixed duration of 0.75 s; (iii) the response window (maximum duration was 2.25 s); (iv) the delay period preceding feedback (3.5–11.5 s, uniformly distributed across 5 levels, steps of 2 s); (v) the feedback and the inter-trial interval (ITI; 3.5–11.5 s, uniformly distributed across 5 levels, steps of 2 s). Stimulus onset coincided with a visual and auditory cue. The auditory cue was presented for 0.25 s (white noise or pure tone at 880 Hz, 50–50% of trials, randomly intermixed). The visual cue was a change in the region of fixation from an open to a closed rectangle. The return of the fixation region to an open rectangle indicated to subjects to give their response (the surface areas in pixels of the open and closed rectangles were held equal in order to assure no change in overall luminance). Feedback was presented visually (green/red for correct/error) for 50 frames (0.42 s at 120 Hz). If subjects did not respond or were too fast/slow in responding, a yellow rectangle was presented as feedback on that trial.

Each block of the task began and ended with a 12-s baseline period, consisting of a fixation region (no dots). Each block of the task had 25 trials and lasted approximately 8 minutes. Subjects performed between 23 and 24 blocks yielding a total of 575–600 trials per subject. One subject performed a total

of 18 blocks (distributed over three sessions), yielding a total of 425 trials. Data from one session of two subjects (12 blocks in total) and 2 blocks of a third subject were excluded from the analyses because of poor eye-tracker data quality or technical error.

4.4.3 Visual stimuli

Dot motion stimuli were presented within a central annulus that was not visible to the subjects (grey background, outer diameter 16.8°, inner diameter of 2.4°). The fixation region was in the center of the annulus and consisted of a black rectangle (0.45°length). Signal dots moved at 7.5°/s in one of two directions (90°or 270°). Noise dots were randomly assigned (uniformly distributed) to locations within the annulus on each frame, preventing them from being trackable. Each frame consisted of 524 white dots (0.15°in diameter) within one visual hemifield (left or right; The hemifield remained constant during a block of trials and was counterbalanced between blocks. This manipulation was specific for the MRI experiment. The two hemifields were averaged in the current analysis). The proportion of 'signal' as compared with 'noise' dots defined motion coherence levels. Signal dots were randomly selected on each frame, lasted 10 frames, and were thereafter re-plotted in random locations (reappearing on the opposite side when their motion extended outside of the annulus). To prevent tracking of individual dots, independent motion sequences (n = 3) were interleaved [106].

4.4.4 Eye-tracking data acquisition and preprocessing

Pupil diameter was measured using an EyeLink 1000 Long Range Mount (SR Research, Ottawa, Ontario, Canada). Either the left or right pupil was tracked (via the mirror attached to the head coil) at 1000 Hz sample rate with an average spatial resolution of 15 to 30 min arc. The MRI681 compatible (non-ferromagnetic) eye tracker was placed outside the scanner bore. Eye position was calibrated once at the start of each scanning session.

Eye blinks and saccades were detected using the manufacturer's standard algorithms (default settings). Further preprocessing steps were carried out using custom-made Python software, which consisted of (i) linear interpolation around blinks (time window from 0.1 s before until 0.1 s after each blink), (ii) band-pass filtering (third-order Butterworth, passband: 0.01–6 Hz), (iii) removing responses to blink and saccade events using multiple linear regression (responses estimated by deconvolution) [108], and (iv) converting to percent signal change with respect to the mean of the pupil time series per block of trials.

4.4.5 Quantifying pre- and post-feedback pupil responses

Pupil dilation is affected by a range of non-cognitive factors [217], whose impact needs to be eliminated before inferring the relation between central arousal and computational variables of interest. We excluded the impact of a number of non-cognitive factors on the pupil responses: (i) blinks and eye movements, which were eliminated from the analysis (see above); (ii) luminance, which was held constant throughout

the trial, with the exception of the visual feedback signals, which we controlled for in a separate control experiment: Supp. Fig. S2); (iii) motor responses [133]; and (iv) trial-by-trial variations in decision time that may confound pupil response amplitudes [33, 40] due to the temporal accumulation properties of the peripheral pupil apparatus [41, 109]. With the aim of excluding effects related to above mentioned points (iii) and (iv), we investigated pupil responses locked to the choice reported by the observer. Additionally, only trials with the three longest delay intervals between events (7.5, 9.5 and 11.5 s; 3/5 of all trials) were used in the main analysis of pupil responses. Specifically, for the pre-feedback interval, the delay period was between the choice and feedback. For the post-feedback interval, the delay period was the inter-trial interval. Finally, we performed a control analysis in which RTs were removed from pupil responses via linear regression (see Fig. 4).

For each trial of the motion discrimination task, two events of interest were inspected: (a) pupil responses locked to the observers' reported choice (button press) and (b) pupil responses locked to the onset of the feedback. On each trial, the mean baseline pupil diameter (the preceding 0.5 s) with respect to the motion stimulus onset and feedback onset was subtracted from the evoked and mean responses for the pre-feedback and post-feedback intervals, respectively. We extracted the mean pupil responses within the sustained time window (3–6 s), defined by the period during which the independently measured pupil IRF returned to baseline (at the group level, Fig. 3a). The uncertainty signal was expected to be largest in the time window just preceding feedback based on Urai et al.3, reflecting the fact that the 'reward anticipation' state is highest the longer the observer waits for feedback. Therefore, we additionally analyzed pre-feedback pupil responses in the 0.5 s preceding feedback.

4.4.6 Model predictions

In signal detection theory, on each trial a decision variable (dv_i) was drawn from a normal distribution $N(\mu, \sigma)$, where μ was the sensory evidence on the current trial and σ was the level of internal noise. In our case, we took μ to range from -0.5 to 0.5, corresponding to the extremes of the motion coherence presented in the main experiment (where 0 = 100% random motion and 1 = 100% coherent motion). The internal noise, σ, was estimated by fitting a probit psychometric function onto the combined data across all subjects (slope β = 7.5). The standard deviation, σ, of the dv distribution is $\frac{1}{\beta} = 0.133$. The decision bound, c, was set to 0, indicating no choice bias for any observer.

For each level of evidence strength, $\mu = [-0.5, 0.5]$ in steps of 0.01, we simulated a normal distribution of dv with σ = 0.133 with 10,000 trials. The choice on each trial corresponded to the sign of dv_i. A choice was correct when the sign of dv_i was equal to the sign of μ_i. Errors occurred due to the presence of noise in the dv, which governed choice in both of the two models discussed as follows.

We simulated two models, the Belief State Model and the Stimulus State Model, which differed only in the input into the function used to compute confidence: whether the confidence is a function of dv_i or μ_i. Confidence was defined as

$$BeliefStateConfidence = \frac{1}{n} \times \sum_{i=1}^{n} f(|dv_i - c|) \tag{1}$$

$$StimulusStateConfidence = \frac{1}{n} \times \sum_{i=1}^{n} f(|\mu_i - c|) \tag{2}$$

where n was the number of trials per condition, for which the predictions were generated (see below), f was the cumulative distribution function of the normal distribution, transforming the distance $|dv-c|$ or $|\mu-c|$ into the probability of a correct response, for the Belief State or Stimulus State Model, respectively.

$$f(x) = \frac{1}{2}[1 + erf(\frac{x}{\sigma\sqrt{2}})] \tag{3}$$

Because we applied equations 1 and 2 separately to each combination of Difficulty (i.e. coherence level) and Accuracy (Error and Correct) conditions, n depended on the variable number of trials obtained in each condition (with the smallest n for the Easy Error condition) in our simulations.

Decision uncertainty was the complement of confidence

$$Uncertainty = 1 - confidence \tag{4}$$

And the prediction error was defined as

$$Predictionerror = feedback - confidence \tag{5}$$

Pre-feedback pupil responses have previously been found to reflect decision uncertainty3; we therefore expected the post-feedback pupil responses to similarly follow the complement of the prediction error (i.e. 1-prediction error). For each trial, we computed the binary choice, the level of decision uncertainty, the accuracy of the choice and the prediction error. For plotting, we collapsed the coherence levels across the signs of μ, as these are symmetric for the up and down motion directions.

Custom Python code used to generate the model predictions can be found here: `https://github.com/colizoli/pupil_belief_states`.

4.4.7 Motion energy

To extract estimates of fluctuating sensory evidence, we applied motion energy filtering to the single-trial dot motion stimuli (using the filters described in [210]). Summing the 3D motion energy values over space and time gave us a single-trial estimate of the external sensory evidence presented to the subject (positive for upwards, negative for downwards motion). We used the absolute value of this signed motion energy signal as our continuous measure of sensory evidence strength in statistical analyses. For visualization (Fig. 4b), we divided this absolute motion energy metric into 4 equally-sized bins within every observer.

4.4.8 Statistical analysis

Behavioral variables and pupil responses were averaged for each condition of interest per subject (N = 15). Statistical analysis of mean differences in pupil dilation of evoked responses was done using cluster-based permutation methods65. The average responses in the sustained time windows were evaluated using a two-way ANOVA with factors: difficulty (2 levels: Hard vs. Easy) and accuracy (2 levels: Correct vs. Error). All post-hoc and two-way comparisons were based on non-parametric permutation tests (two-tailed).

4.4.9 Control experiment 1: Individual pupil impulse response functions

In order to define a sustained component of pupil responses evoked by the events of interest during the main experiment, we independently measured subjects' pupil responses evoked by simply pushing a button upon hearing a salient cue. This enabled a principled definition of the time window of interest in which to average pupil responses based on independent data. Subjects performed one block of the pupil impulse response task at the start of each experimental session (while anatomical scans were being acquired). Pupil responses following an auditory cue were measured for each subject63. Pupils were tracked while subjects maintained fixation at a central region consisting of a black open rectangle (0.45°length) against a grey screen. No visual stimuli changed, ensuring constant illumination within a block. An auditory white noise stimulus (0.25 s) was presented at random intervals between 2 and 6 s (drawn from a uniform distribution). Participants were instructed to press a button with their right index finger as fast as possible after each auditory stimulus. One block consisted of 25 trials and lasted 2 min. Two subjects performed three blocks, yielding a total of 75–100 trials per subject. Trials without a response were excluded from the analysis. Each subject's impulse response function (IRF) was estimated using deconvolution (with respect to the auditory cue) in order to remove effects of overlapping events due to the short delay interval between subsequent trials61.

4.4.10 Control experiment 2: Pupil responses during passive viewing of feedback signals

Pupil responses evoked by the green and red fixation regions used in the main experiment were measured in a separate control experiment (see Supplementary Fig. S2; N = 15, 5 women, aged 28.5 ± 4 years, range 23–34). Three subjects were authors, two of which participated in the main 2AFC task. No other subjects from this control experiment participated in the main 2AFC task. Pupils were tracked while subjects maintained fixation at a central region of the screen. Stimuli were identical to the main 2AFC task; dot motion consisted of only random motion (0% coherence). A trial consisted of two phases: (i) the baseline period preceding the onset of a color change (1–3 s, uniform distribution), and (ii) passive viewing of the stimuli used for feedback in the main experiment: during which the fixation region changed to either red or green (50–50% of trials, randomized) for 50 frames (0.42 s at 120 Hz). This was followed by an ITI

(3–6 s, uniformly distributed). Participants were instructed that they did not need to respond, only to maintain fixation. A block consisted of 25 trials and lasted 3 min. Subjects performed eight blocks of this task in the behavioral lab, yielding 200 trials per subject.

4.4.ii Data availability

The pupil data and model prediction code are publicly available here: https://github.com/colizoli/pupil_belief_states.

4.5 Acknowledgements

This research was supported by the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 604102 (Human Brain Project) (to T.H.D.), the German Research Foundation (DFG): DO 1240/3-1, DO 1240/2-1, and SFB 936/A7 (to T.H.D.) and the German Academic Exchange Service (DAAD) (to A.E.U.). This work was carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative (www.surfsara.nl). We thank Tomas Knapen for contributing analysis materials.

4.6 Supplementary Figures

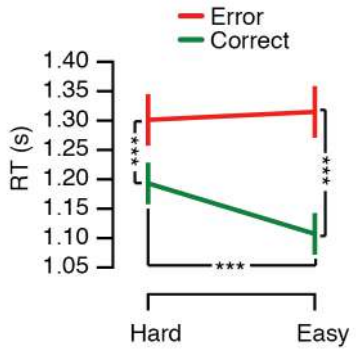


Figure S1: RT scales with decision uncertainty. Mean reaction times (RT) as a function of task difficulty and accuracy. Task difficulty and accuracy interacted. Error bars represent the standard error of the mean (N = 15). ***p < 0.001.

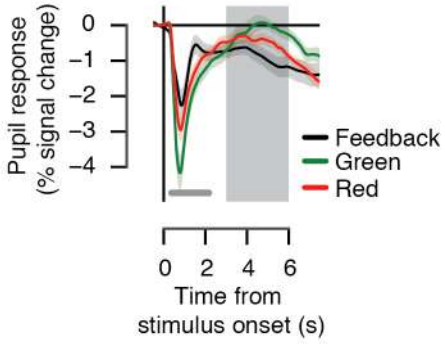


Figure S2: Pupil responses during passive viewing of feedback signals. In a control experiment (N = 15, 5 women, aged 28.5±4 years, range 23-34), we investigated the time course of potential differences in pupil responses evoked by red as compared with green light, regardless of whether these colors correspond to reward feedback during the perceptual choice task. Three subjects were authors, two of which participated in the main experiment. Stimuli were identical to the main 2AFC task; dot motion consisted of only random motion (0% coherence). A trial consisted of a baseline period preceding the onset of a color change (1-3 s, uniformly distributed), the red or green rectangle at fixation (50-50% of trials, randomized, 0.42 s), and ITI (3-6 s, uniformly distributed). Participants passively viewed the stimuli while maintaining fixation. Pupil responses were averaged for each condition of interest per subject (N = 15, 200 trials per subject). The light-mediated pupil constrictions evoked by visual feedback cues during the main task (grey) and in the passive viewing control experiment (red, green). The grey bar indicates a difference between red- and green-evoked responses, p < 0.05 (cluster-based permutation test, see main text). Grey shaded area, 'sustained' time window during which pupil dilation was averaged, defined by the period during which the pupil impulse response function returned to baseline and the shortest delay between events (3-6 s). The results show that (i) green and red light both evoked pupil constrictions, and (ii) green light produced slightly larger pupil constriction than red light, in an early time window (0.25-2.25 s). The difference in Correct vs. Error trials in pupil constriction after feedback during the main experiment continues after this early time window (see Figure 3c). Furthermore, any differences obtained within Error and Correct conditions after feedback during the main experiment cannot be explained by differences between the color-evoked responses, as the stimulus color was the same between these comparisons.

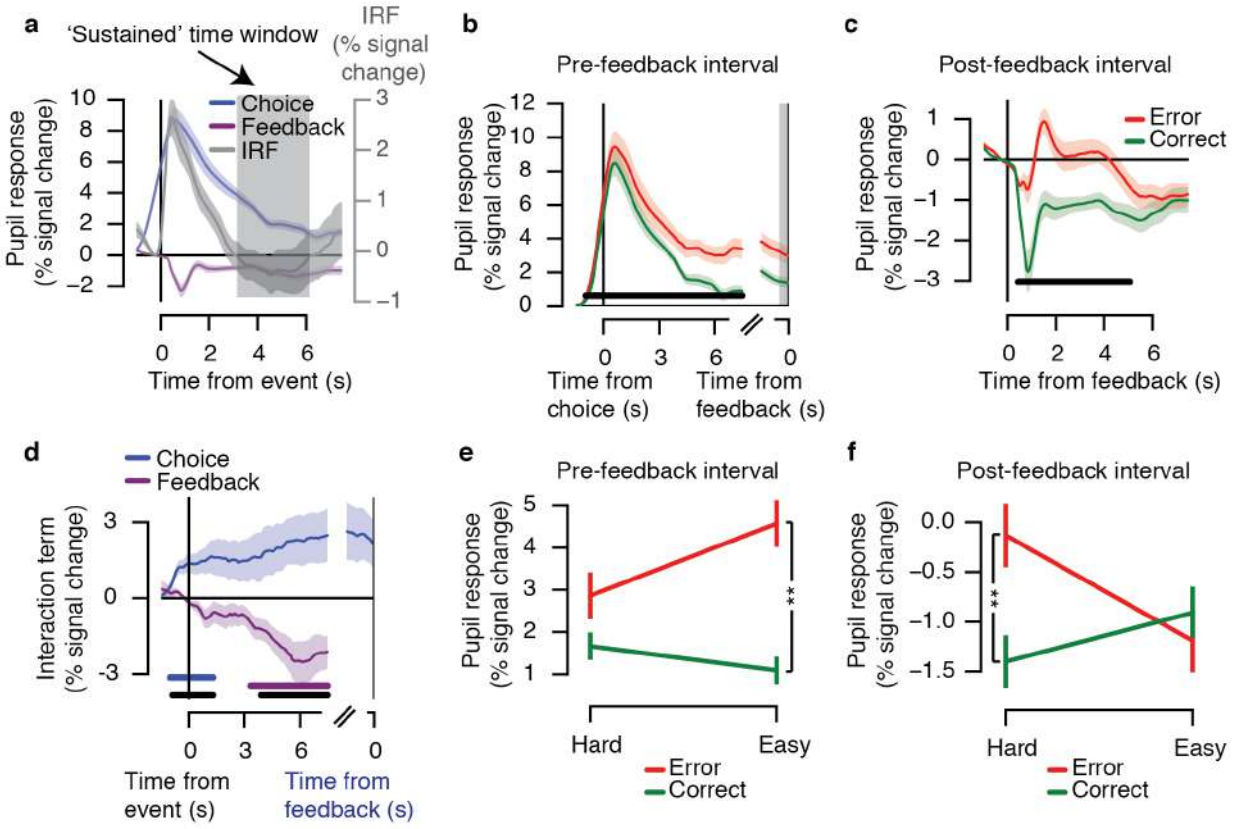


Figure S3: Replication of Figure 3 with all trials. The same pattern of pupil responses was obtained when all trials were included in the analysis, including those in which the interval between events (Delay and ITI periods of a trial; see Fig. 1) was less than 7.5 s (delays ranged from 3.5-11.5 s, with 5 levels in steps of 2 s, uniformly distributed) **(a-f)**. For the pupil responses in the -0.5 s window preceding feedback **(e)**, a significant interaction between difficulty and accuracy was obtained in this later time window ($F(1,14) = 4.95$, $p = 0.043$; post hoc comparisons: Hard Error vs. Hard Correct, $p = 0.072$; Easy Error vs. Easy Correct, $p = 0.004$; Hard Error vs. Easy Error, $p = 0.074$; Hard Correct vs. Easy Correct, $p = 0.061$). During the post-feedback interval, a significant interaction between difficulty and accuracy was obtained **(f)**, $F(1,14) = 7.89$, $p = 0.014$; post hoc comparisons: Hard Error vs. Hard Correct, $p = 0.001$; Easy Error vs. Easy Correct, $p = 0.572$; Hard Error vs. Easy Error, $p = 0.066$; Hard Correct vs. Easy Correct, $p = 0.118$).

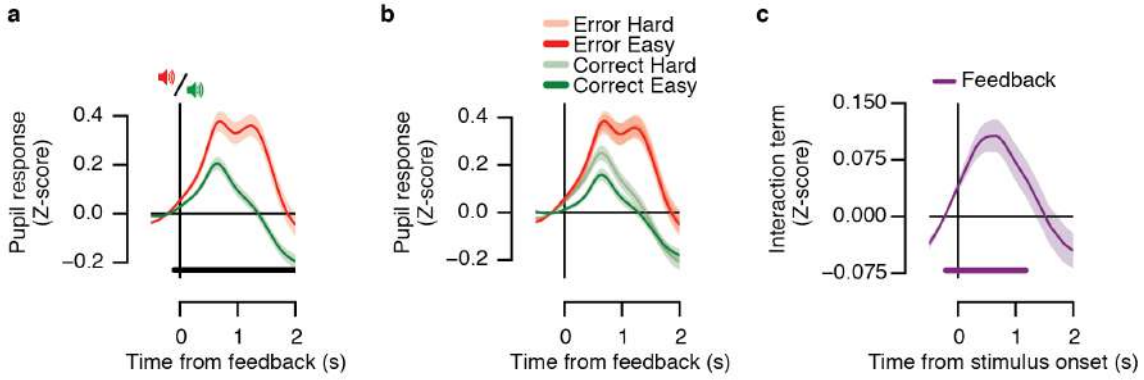


Figure S4: Reanalysis of the data from our previously published study [125] (available at https://doi.org/10.6084/m9.figshare.4300043). This study used a similar visual perceptual choice task, however with a number of important differences specified in the following: The study used a two-interval forced choice motion coherence discrimination task; multiple levels of task difficulty were intermixed, here sorted into two categories (median split) yielding Hard and Easy conditions for comparison with the present data; delay intervals between decision and feedback, and the inter-trial-intervals were shorter than in the current study; feedback (Correct or Error) was presented by two different tones; feedback was not linked to any reward (participants' financial remuneration was not contingent on performance). **(a)** Evoked pupil responses for Correct (green) and Error (red) trials locked to trial-wise (auditory) feedback. The black bar indicates Correct vs. Error effect, $p < 0.05$ (cluster-based permutation test). Because feedback was not presented visually, there was no post-feedback pupil constriction, but dilation for all trial types. Error feedback elicited stronger dilations than correct feedback, as in the current data (compare with Figure 3c). **(b)** Pupil responses as a function of task difficulty and accuracy locked to feedback. The scaling with evidence strength was similar to pre-feedback decision uncertainty, but not to post-feedback prediction error, with smaller dilations for Correct Easy than Correct Hard responses (compare to Figure 2c). **(c)** The interaction term for task difficulty with two levels (Easy Error - Easy Correct) - (Hard Error - Hard Correct) for feedback-locked responses. The purple bar indicates the feedback-locked response tested against 0, $p < 0.05$ (cluster-based permutation test). For all feedback-locked responses, the mean pupil diameter across the pre-feedback interval from -0.5 s to 0 s was subtracted from the response time courses at the single-trial level. Each condition of interest was averaged across subjects (N = 27).

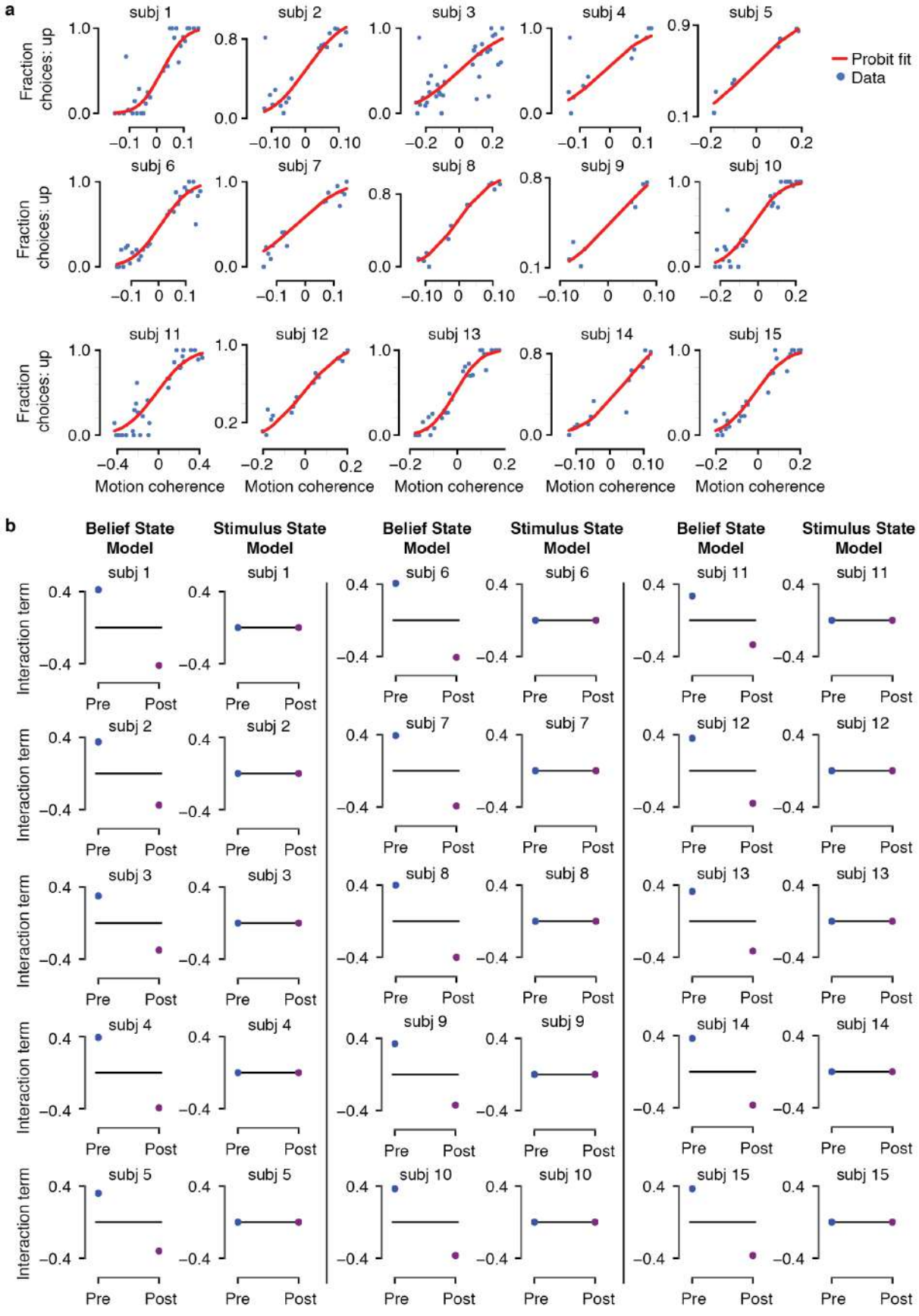

a
subj 1
subj 2
subj 3
subj 4
subj 5
subj 6
subj 7
subj 8
subj 9
subj 10
subj 11
subj 12
subj 13
subj 14
subj 15
Fraction choices: up
Motion coherence
Probit fit
Data
b
Belief State Model
Stimulus State Model
Interaction term
Pre
Post

Figure S5: Individual psychometric functions and model predictions. Based on the individual estimates of internal noise in the data (i.e. sigma) **(a)**, subject-specific model predictions were generated for the Belief State and Stimulus State models **(b)**. Predictions for the interaction term defined as (Easy Error - Easy Correct) - (Hard Error - Hard Correct) based on subject-specific motion coherence levels are shown.

Status: published
Urai AE, de Gee JW, Tsetsos K & Donner TH (2019). *eLife,* 8, 8, e46331.

5

Choice history biases subsequent evidence accumulation

Conceptualization by AEU, THD; Investigation by AEU, JWG; Formal analysis by AEU, JWG, KT; Resources by THD; Writing–original draft by AEU, THD; Writing–review and editing by AEU, JWG, KT, THD.

Perceptual choices not only depend on the current sensory input, but also on the behavioral context, such as the history of one's own choices. Yet, it remains unknown how such history signals shape the dynamics of later decision formation. In models of decision formation, it is commonly assumed that choice history shifts the starting point of accumulation towards the bound reflecting the previous choice. We here present results that challenge this idea. We fit bounded-accumulation decision models to human perceptual choice data, and estimated bias parameters that depended on observers' previous choices. Across multiple task protocols and sensory modalities, individual history biases in overt behavior were consistently explained by a history-dependent change in the evidence accumulation, rather than in its starting point. Choice history signals thus seem to bias the interpretation of current sensory input, akin to shifting endogenous attention towards (or away from) the previously selected interpretation.

5.1 Introduction

Decisions are not isolated events, but are embedded in a sequence of choices. Choices, or their outcomes (e.g. rewards), exert a large influence on subsequent choices [12, 13]. This holds even for low-level perceptual choices [14–16]. In most perceptual choice tasks used in the laboratory, the decision should only be based on current sensory input, the momentary "evidence" for the decision. Thus, most work on their computational and neurophysiological mechanisms has largely focused on the transformation of sensory evidence into choice [20]. Yet, perceptual decisions are strongly influenced by a number of historical factors: whether or not previous choices led to positive outcomes [15, 225], the confidence in them [226], and their content (i.e. which stimulus categories were selected; [125, 227, 228]. The latter type of sequential effect, which we call "choice history bias", refers to the selective tendency to repeat (or alternate) previous choices. It is distinct and dissociable from effects of reward, performance feedback or subjective error awareness in previous trials.

Choice history biases are prevalent in human [125, 228], monkey [229] and rodent [230, 231] perceptual decision-making. Remarkably, this holds even for environments lacking any correlations between stimuli presented on successive trials – the standard in psychophysical laboratory experiments. Choice history biases vary substantially across individuals [125, 232]. Neural signals reflecting previous choices have been found across the sensorimotor pathways of the cerebral cortex, from sensory to associative and motor regions [85, 227, 229, 233–238].

By which mechanism are choice history signals incorporated into the formation of a decision? Current models of perceptual decision-making posit the temporal accumulation of sensory evidence, resulting in an internal decision variable that grows with time [2, 18, 19, 25]. When this decision variable reaches one of two decision bounds, a choice is made and the corresponding motor response is initiated. In this framework, a bias can arise in two ways: (i) by shifting the starting point of accumulation towards one of the bounds, or (ii) by selectively changing the rate at which evidence for one versus the other choice alternative is accumulated. Figure 1 illustrates these two biasing mechanisms for a simple and widely used form of accumulation-to-bound model: the drift diffusion model (DDM). Similar principles apply to more complex accumulation-to-bound models we will explore below. The starting point shift can be thought of as adding an offset to the perceptual interpretation of the current sensory evidence. By contrast, the evidence accumulation bias corresponds to biasing that perceptual interpretation towards one of the two stimulus categories.

It is currently unknown which of those two principal mechanisms accounts for the choice history biases observed in overt behavior. Previous theoretical accounts have postulated a shift in the starting point of the decision variables towards the bound of the previous choice [239–241]. This is based on the assumption that the representation of the decision variable decays slowly, leaving a trace of the observer's choice in the next trial [242–245]. However, choice history biases might also originate from a slower (i.e., tens of seconds) across-trial accumulation of internal decision variables – analogous to the accumulation of external outcomes in value-based decisions [12, 13]. Previous experimental work on perceptual choice

history biases either did not analyze the within-trial decision dynamics [85, 125, 208, 227, 228, 230], or only tested for starting point biases, not accumulation biases [229, 239, 240, 242, 243, 246–248].

Here, we untangle how history-dependent changes in evidence accumulation and starting point contribute to history biases in overt choice behavior. Across a range of perceptual choice tasks, we found that individual differences in choice repetition are explained by history-dependent biases in accumulation, not starting point. Thus, the interaction between choice history and decision formation seems to be more complex than previously thought: choices may bias later evidence accumulation processes towards (or away from) the previous chosen perceptual interpretation of the sensory input.

5.2 Results

We fit different bounded-accumulation models to human behavioral data (choices and response times (RT). The DDM fits model parameters from joint choices RT distributions, and provides good fits to behavioral data from a large array of two-choice task [25]. We estimated the following parameters: non-decision time (the time needed for sensory encoding and response execution), starting point of the decision variable, separation of the decision bounds, mean drift rate, and a stimulus-independent constant added to the mean drift. We refer to the latter parameter (termed "drift criterion" by [25]) as "drift bias".

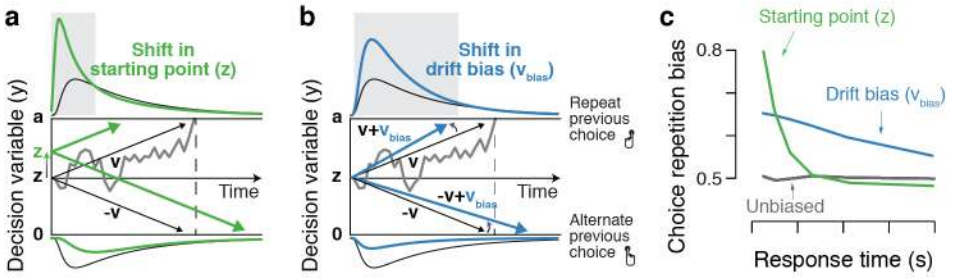


Figure 1: Two biasing mechanisms within the DDM. The DDM postulates that noisy sensory evidence is accumulated over time, until the resulting decision variable y reaches one of two bounds (dashed black lines at y=0 and y=a) for the two choice options. Repeating this process over many trials yields RT distributions for both choices (plotted above and below the bounds). Gray line: example trajectory of decision variable from single trial. Black lines: mean drift and resulting RT distributions under unbiased conditions. **(a)** Choice history-dependent shift in starting point. Green lines: mean drift and RT distributions under biased starting point. Grey shaded area indicates those RTs for which starting point leads to choice bias. **(b)** Choice history-dependent shift in drift bias. Blue lines: mean drift and RT distributions under biased drift. Grey shaded area indicates those RTs for which drift bias leads to choice bias. **(c)** Both mechanisms differentially affect the shape of RT distributions. Conditional bias functions [249], showing the fraction of biased choices as a function of RT, demonstrate the differential effect of starting point and drift bias shift.

Within the DDM, choice behavior can be selectively biased toward repetition or alternation by two mechanisms: shifting the starting point or biasing the drift towards (or away from) the bound for the previous choice (Figure 1). These biasing mechanisms are hard to differentiate based on the proportion of choices alone, but they are readily distinguishable by the relationship between choice bias and RT (Figure 1c). Specifically, the conditional bias function [249] shows the fraction of choices that are repetitions as a function of their RT (binned in quantiles). A shift in starting point is most influential early on in the decision process: it affects the leading edge of the RT distribution and shifts its mode. It predicts that the

majority of history-dependent choice biases occur on trials with fast RTs (Figure 1c, green). A drift bias is instead accumulated along with the evidence and therefore grows as a function of elapsed time. Thus, drift bias strongly affects the trailing edge of the RT distribution with only a minor effect on the mode, altering choice fractions across the whole range of RTs (Figure 1c, blue). History-dependent changes in bound separation or mean drift rate may also occur, but they can only change overall RT and accuracy: those changes are by themselves not sufficient to bias the accumulation process toward one or the other bound, and thus towards choice repetition or alternation (see Figure 4 – Figure Supplement 4).

We fit different variants of the DDM (Figure 3 – Figure Supplement 1) to data from six different experiments. These covered a range of task protocols and sensory modalities commonly used in studies of perceptual decision-making (Figure 2a and Methods section Datasets: behavioral tasks and participants): two alternative forced-choice, two interval forced-choice, and yes-no (simple forced choice) tasks; RT and so-called fixed duration tasks; visual motion direction and coherence discrimination, visual contrast and auditory detection; and experiments with and without single-trial performance feedback. As found in previous work [125, 228, 232], observers exhibited a wide range of idiosyncratic choice history biases across all experiments (Figure 2b,c). To ensure that the DDM is an appropriate (simplified) modeling framework for these data, we first fit a basic version of the DDM that contained the above-described parameters, without allowing bias parameters to vary with choice history. We then fit the DDM while also allowing starting point, drift bias, or both to vary as a function of the observer's choice on the previous trial.

The DDM fits matched several aspects of the behavioral data well (Figure 3 – Figure Supplement 1). First, RT distributions matched the model predictions reasonably well (shown separately for each combination of stimuli and choices in Figure 3 – Figure Supplement 1, darker colors indicate predicted RTs obtained through model simulations). Second, for the fits obtained with a hierarchical Bayesian fitting procedure (see Figure 3 – Figure Supplement 1 and Materials and Methods), used for Figures 3-5, the $\hat{R}$ for group-level parameters ranged between 0.9997 and 1.0406 across datasets, indicating good convergence of the sampling procedure. Third, individual drift rate estimates correlated with individual perceptual sensitivity (d', Figure 3 – Figure Supplement 1a) and monotonically increased with stronger sensory evidence (Figure 3 – Figure Supplement 1a). In fixed duration tasks, the decision-maker does not need to set a bound for terminating the decision [18], so the bounded diffusion process described by the DDM might seem inappropriate. Yet, the success of the DDM in fitting these data was consistent with previous work [247, 250, 251] and might have reflected the fact that observers set implicit decision bounds also when they do not control the stimulus duration [252] but see [253].

5.2.1 History-dependent accumulation bias, not starting point bias, explains individual differences in choice repetition behavior

Models with history-dependent biases better explained the data than the baseline model without such history dependence (Figure 3a), corroborating the observation that observers' behavior showed consider-

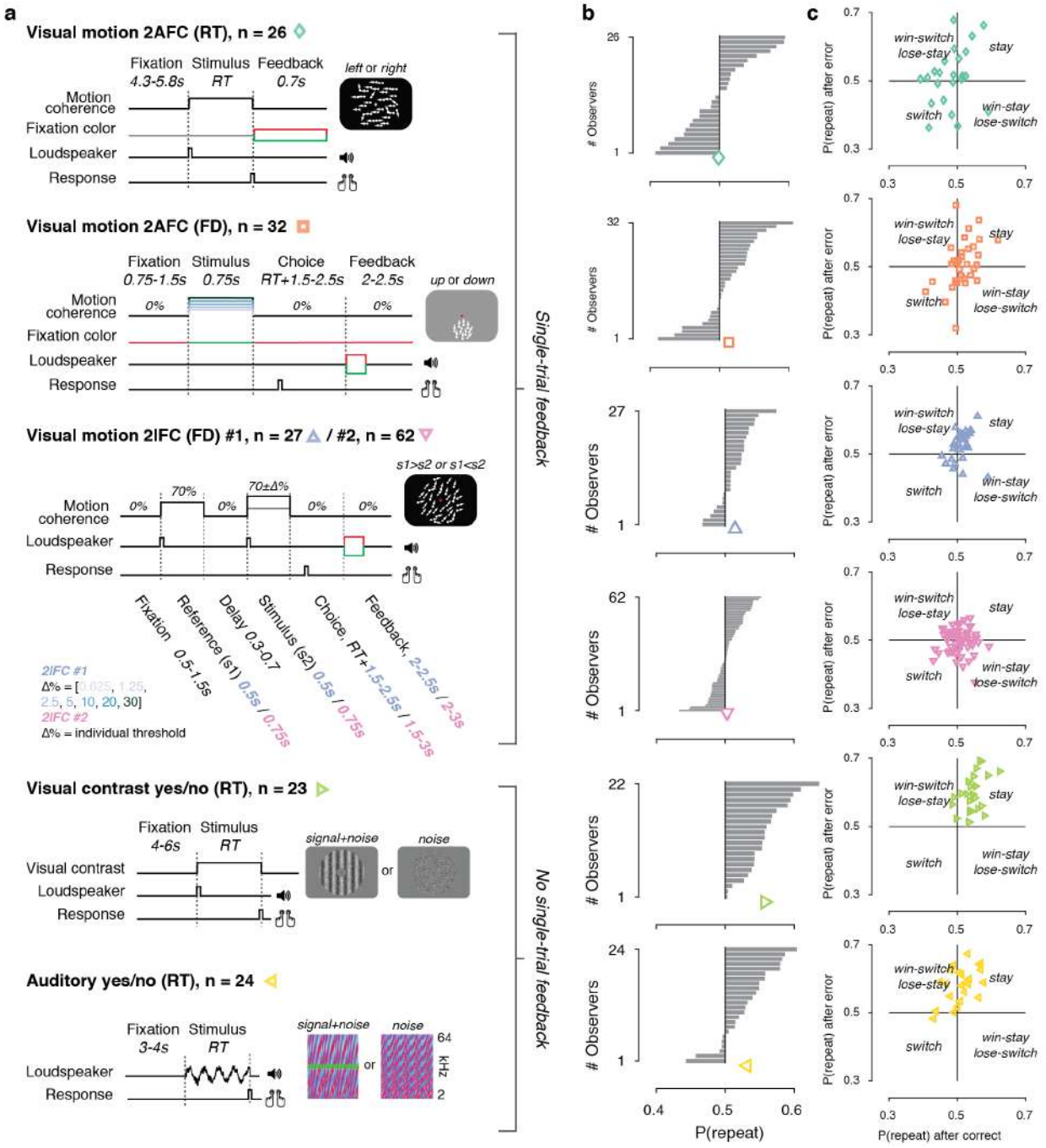


Figure 2: Behavioral tasks and individual differences. **(a)** Schematics of perceptual decision-making tasks used in each dataset. See Methods section *Datasets: behavioral tasks and participants*. **(b)** Distributions of individual choice history biases for each dataset. Grey bars show individual observers, with colored markers indicating the group mean. **(c)** Each individual's tendency to repeat their choices after correct vs. error trials. The position of each observer in this space reflects their choice- and outcome-dependent behavioral strategy.

able dependence on previous choices (Figure 2f). The model with both history-dependent starting point and drift bias provided the best fit to five out of datasets (Figure 3a), based on the Akaike Information Criterion (AIC; [254] – note that we obtained the same results when instead using the hierarchical Deviance

Information Criterion).

The above model comparison pointed to the importance of including a history-dependency into the model. We further examined the ability of each model to explain specific diagnostic features in the data [255] that distinguished starting point from drift bias. A history-dependent shift in the starting point leads to biased choices primarily when responses are fast (early RT quantiles), whereas a history-dependent shift in drift leads to biased choices across all trials, including those with slow responses (Figure 1). We simulated choices and RTs from the three different model variants and computed so-called 'conditional bias functions' [249]: the fraction of choices in line with each observer's choice repetition tendency (i.e., repetition probability) within each quantile of their RT distribution. For observers whose choice repetition probability was > 0.5, this was the fraction of repetitions; for the other observers, this was the fraction of alternations. Consistent with a shift in drift bias, observers exhibited history-dependent choice biases across the entire range of RTs (Figure 3b). In particular, the biased choices on slow RTs could only be captured by models that included a history-dependent shift in drift bias (Figure 3c, blue and dark green bars).

We used the parameter estimates obtained from the full model (with both history-dependent starting point and drift bias) to investigate how history-dependent variations in starting point and drift bias related to each individual's tendency to repeat their previous choices. We call each bias parameter's dependence on the previous choice its 'history shift'. For instance, in the left vs. right motion discrimination task, the history shift in starting point was computed as the difference between the starting point estimate for previous 'left' and previous 'right' choices, irrespective of the category of the current stimulus. The history shift in drift bias, but not the history shift in starting point, was robustly correlated to the individual probability of choice repetition (Figure 4a, significant correlations indicated with solid regression lines). In five out of six datasets, the correlation with the history shift in drift bias was significantly stronger than the correlation with the history shift in starting point (Figure 4b, $\Delta\rho$ values).

We quantified the total evidence by computing a Bayes factor for each correlation [256], and multiplying these across datasets [257]. This further confirmed that individual choice history biases were not captured by history shifts in starting point, but consistently captured by history shifts in drift (Figure 4b). Specifically, the Bayes factor for the history shift in starting point approached zero, indicating strong evidence for the null hypothesis of no correlation. The Bayes factor for the history shift in drift indicated strong evidence for a correlation [258].

Correlations between estimated history shifts in starting point and drift bias were generally negative (mean Spearman's ρ: -0.2884, range -0.4130 to -0.0757), but reached statistical significance ($p < 0.05$) in only one dataset. The combined Bayes Factor BF_{10} was 0.0473, indicating strong evidence for H0. We thus remain agnostic about the relationship between the history shifts of both parameters.

The same qualitative pattern of results was obtained with an alternative fitting procedure (non-hierarchical G2 optimization, Figure 4 – Figure Supplement 1a), as well as a model that allowed for additional across-trial variability in non-decision time (Figure 4 – Figure Supplement 1b). Letting non-decision time vary

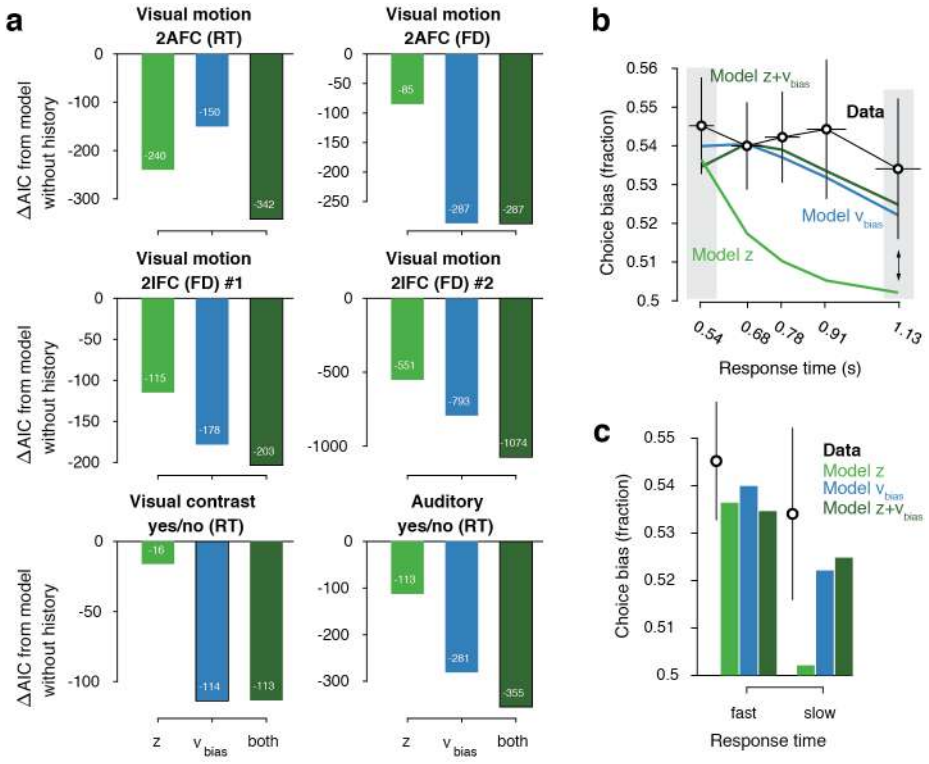


Figure 3: Model comparison and simulations. (a) For each dataset, we compared the AIC between models where drift bias, starting point bias or both were allowed to vary as a function of previous choice. The AIC for a model without history dependence was used as a baseline for each dataset. Lower AIC values indicate a model that is better able to explain the data, taking into account the model complexity; a Δ AIC of 10 is generally taken as a threshold for considering one model a sufficiently better fit. **(b)** Conditional bias functions (Figure 1c; White and Poldrack, 2014). For the history-dependent starting point, drift bias and hybrid models, as well as the observed data, we divided all trials into five quantiles of the RT distribution. Within each quantile, the fraction of choices in the direction of an individual's history bias (repetition or alternation) indicates the degree of choice history bias. Error bars indicate mean ± s.e.m. across datasets. **(c)** Choice bias on slow response trials can be captured only by models that include history-dependent drift bias. Black error bars indicate mean ± s.e.m. across datasets, bars indicate the predicted fraction of choices in the first and last RT quantiles.

with each level of sensory evidence strength (in the two datasets including multiple such levels) did not change the pattern of model comparison and correlation results (Figure 4 – Figure Supplement 1c). These findings are thus robust to specifics of the model and fitting method. The Visual motion 2IFC #2 also included pharmacological interventions in two sub-groups of participants (see Methods); we found the same effects for both drug groups as well as the placebo group (Figure 4 – Figure Supplement 2). A significant positive correlation between history shift in drift bias and P(repeat) was evident separately for two sub-groups of participants, defined by splitting the groups into "Repeaters" and "Alternators" based on P(repeat) being larger or smaller than 0.5, respectively (Figure 4 – Figure Supplement 3).

The lack of a correlation between history-dependent starting point shifts and individual choice repetition is surprising in light of previous accounts [239, 243]. History shifts in starting point were mostly negative (i.e., tendency towards choice alternation) across participants, regardless of their individual tendency towards choice repetition or alternation (Figure 4 – Figure Supplement 5, significant in two out of six datasets). This small but consistent effect likely explains why our formal model comparison favored a

model with both history-dependent drift and starting point over one with drift bias only (see also Discussion). Critically, only the history-dependent shift in drift accounted for individual differences in choice repetition (Figure 4).

5.2.2 History-dependent accumulation bias explains individual choice repetition behavior irrespective of previous choice outcome

In four out of six tasks, participants received explicit outcome feedback (correct, error) after each choice. It is possible that participants experienced positive feedback as rewarding and (erroneously) assumed that a rewarded choice is more likely to be rewarded on the next trial. Manipulations of reward (probability or magnitude) have been found to change starting point [259–261], but might also bias drift [262–264]. Given that there were far more correct (i.e. rewarded) choices than errors, the history-dependent drift bias could reflect the expectation of reward for the choice that was correct on the previous trial.

Two findings refute this idea. First, the same result holds in the two datasets without single-trial outcome feedback (Figure 4a, bottom row), implying that external feedback is not necessary for history shifts in drift bias. Second, we found similar results when separately estimating the model parameters (history shift in starting point and drift bias) and model-free measures (choice repetition probability) after both correct and error trials (Figure 5a). Across datasets, individual repetition probability was best explained by history shifts in drift bias, not starting point, after both correct (Figure 5b) and error (Figure 5c) trials. Thus, even erroneous choices bias evidence accumulation on the next trial, in the same direction as correct choices. Indeed, most participants were predominantly biased by their previous choice (95 'stay', 30 'switch'), while a third was biased by a combination of the previous choice and its correctness (26 'win-stay lose-switch', 42 'win-switch lose-stay'; Figure 2c).

Correlations tended to be smaller for previous erroneous choices. However, directly comparing the correlation coefficients between post-correct and post-error trials (after subsampling the former to ensure equal trial numbers per participant) did not allow us to refute nor confirm a difference (Figure 5d). In sum, history-dependent drift biases did not require external feedback about choice outcome and were predominantly induced by the previous choice. These choice history-dependent biases in evidence accumulation were accompanied by effects on drift rate and boundary separation (Figure 4 – Figure Supplement 4), in line with previous work on post-error slowing [225, 234, 265].

5.2.3 Accumulation bias correlates with several past choices

Does the history shift in evidence accumulation depend on events from one past trial only? Recent work has exposed long-lasting choice history biases that span several trials and tens of seconds [125, 208, 266]. We thus estimated the influence of past events on the evidence accumulation process in a more comprehensive fashion. We fit a family of models in which correct and incorrect choices from up to six previous trials were used as predictors, and estimated their contribution to current starting point and drift bias (see Materials and Methods).

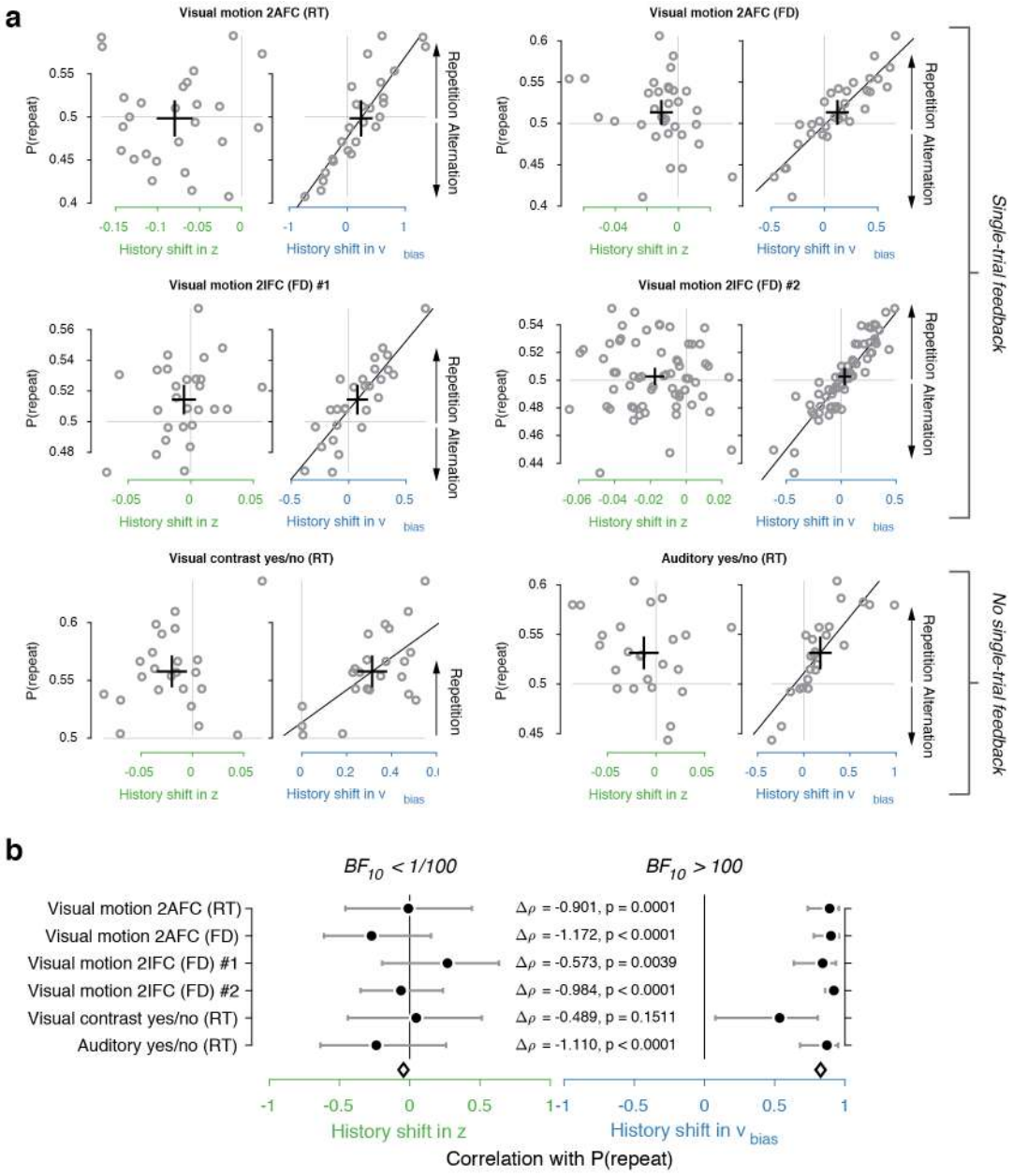


Figure 4: Individual choice history biases are explained by history-dependent changes in drift bias, not starting point. (a) Relationship between individual choice repetition probabilities, P(repeat), and history shift in starting point (left column, green) and drift (right column, blue). Parameter estimates were obtained from a model in which both bias terms were allowed to vary with previous choice. Horizontal and vertical lines, unbiased references. Thick black crosses, group mean ± s.e.m. in both directions. Black lines: best fit of an orthogonal regression (only plotted for correlations significant at $p < 0.05$). **(b)** Summary of the correlations (Spearman's r) between individual choice repetition probability and the history shifts in starting point (green; left) and drift bias (blue; right). Error bars indicate the 95% confidence interval of the correlation coefficient. $\Delta\rho$ quantifies the degree to which the two DDM parameters are differentially able to predict individual choice repetition, p-values from Steiger's test. The black diamond indicates the mean correlation coefficient across datasets. The Bayes factor (BF_{10}) quantifies the relative evidence for the alternative over the null hypothesis, with values < 1 indicating evidence for the null hypothesis of no correlation, and > 1 indicating evidence for a correlation.

Inclusion of further lags improved the model's ability to account for the data, up to a lag of 2-4 after

which model fits (Δ AIC) began to deteriorate (Figure 6 – Figure Supplement 1). In 4/6 datasets, the best-fitting model contained only history-dependent changes in drift, not starting point, over a scale of the previous 2-4 trials. In the other two datasets, the best-fitting model was a hybrid where both drift and starting point varied as a function of choice history, up to 2 to trials into the past (Figure 6 – Figure Supplement 1). We computed "history kernels" across the different lags, separately for starting point and drift bias. These are analogous to the kernels obtained from a history-dependent regression analysis of the psychometric function that ignores decision time [228], and which have been widely used in the recent literature on choice history biases [125, 208, 228]. To interpret these group-level kernels in light of substantial individual variability, we expressed each regression weight with respect to individual repetition probability at lag 1 (i.e. switching the sign for alternators; see Materials and Methods).

Previous choices shifted drift bias in line with individual history bias, across several trials whereas starting point does not consistently shift in the direction of history bias. The hybrid models showed that the effect of choice history on drift bias decayed over approximately 3 past trials (Figure 6a), with a slower decay than for starting point (Figure 6a). The regression weights for past trials (from lag 2 through each dataset's best-fitting lag) for drift bias – but not starting point – significantly correlated with the probability of repeating past choices at these same lags (Figure 6b). This was true after both correct and error trials (Figure 6b), similarly to the effects at lag 1 (Figure 5b-c).

In sum, the biasing effect of choice history on evidence accumulation is long-lasting (longer than the effects on starting point), dependent on preceding choices several trials into the past, but independent of their correctness. This analysis corroborates the previous findings from our simpler models focusing on only the preceding trial and further dissociate the effects of choice history on starting point and evidence accumulation.

5.2.4 History-dependent accumulation bias explains individual choice repetition behavior irrespective of specifics of bounded-accumulation models

We next set out to test the generality of our conclusions and gain deeper mechanistic insight into the nature of the dynamic (i.e., time-increasing) bias. We used a variety of bounded-accumulation models with more complex dynamics than the standard DDM. We focused on the preceding trial only, which our previous analyses had identified as exerting the same effect on history bias as the longer lags (Figure 6). These models included variants of the DDM (i.e. a perfect accumulator) with more complex dynamics of the bias or the decision bounds, as well as variants of a leaky accumulator [53, 68, 267]. We focused on the Visual motion 2AFC (FD) dataset because it entailed small random dot stimuli (diameter 5° of visual angle), leading to large within- and across-trial fluctuations in the sensory evidence which we estimated through motion energy filtering [210, 268]; Supplementary Figure 7a,b). These fluctuating motion energy estimates were used as time-varying sensory input to the models, providing key additional constraints over and above nominal sensory evidence levels, choices and RT distributions [68].

We first re-fit the standard DDM with the two biasing parameters allowed to vary with previous choice

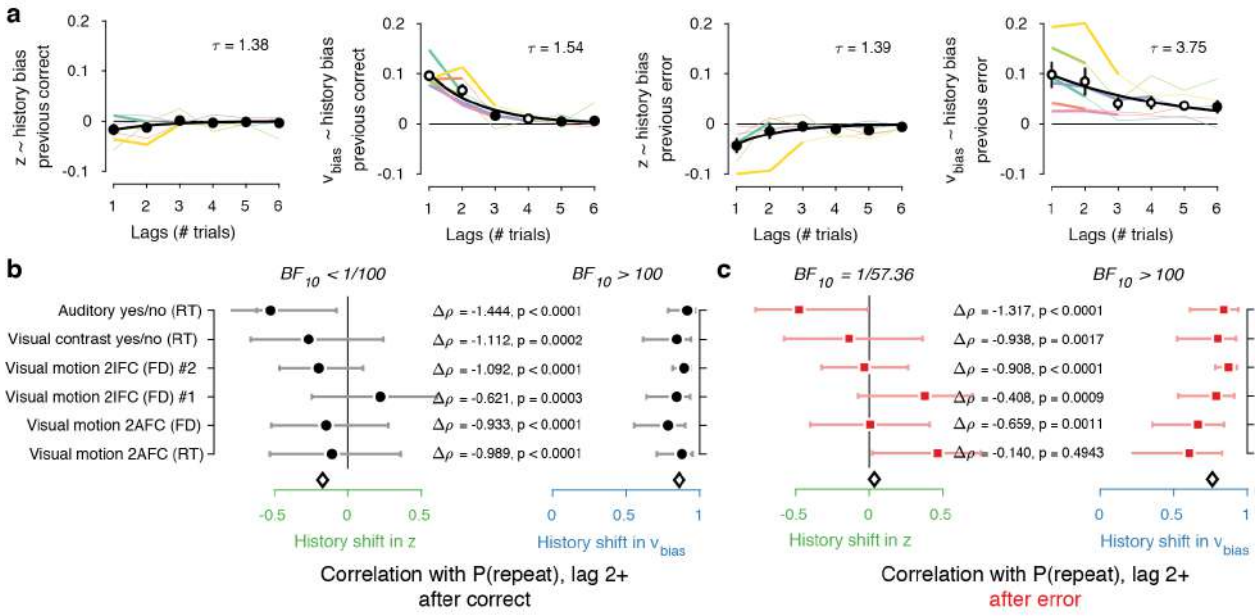


Figure 6: Choice history affects drift bias over multiple trials. (a) History kernels, indicating different parameters' tendency to go in the direction of each individual's history bias (i.e. sign-flipping the parameter estimates for observers with P(repeat) < 0.5). For each dataset, regression weights from the best-fitting model (lowest AIC, Figure 6 – Figure Supplement 1) are shown in thicker lines; thin lines show the weights from the largest model we fit. Black errorbars show the mean +- s.e.m. across models, with white markers indicating timepoints at which the weights are significantly different from zero across datasets ($p < 0.05$, FDR corrected. Black lines show an exponential fit $V(t) = Ae^{-t/\tau}$ to the average data. **(b)** Correlations between individual P(repeat) and regression weights, as in Figure 5b-c. Regression weights for the history shift in starting point and drift bias were averaged from lag 2 until each dataset's best-fitting lag. P(repeat) was corrected for expected repetition at longer lags given individual repetition, and averaged from lag 2 to each dataset's best-fitting lag. $\Delta\rho$ quantifies the degree to which the two DDM parameters are differentially able to predict individual choice repetition, p-values from Steiger's test. The black diamond indicates the mean correlation coefficient across datasets. The Bayes factor (BF_{10}) quantifies the relative evidence for the alternative over the null hypothesis, with values < 1 indicating evidence for the null hypothesis of no correlation, and > 1 indicating evidence for a correlation.

(see Figure 1), now using single-trial motion energy estimates and a non-hierarchical fitting procedure (see Methods). This made these fits directly comparable to both the hierarchical fits in Figures 3-4, and the more complex models described below. As expected (Figure 3a), the data were better explained by a history-dependent bias in the drift, rather than the starting point (Figure 7b1). In these non-hierarchical fits, the hybrid DDM (i.e. both bias terms free to vary as a function of previous choice) lost against the drift bias-only model (indicated by its higher AIC). Yet the hybrid model allowed for a direct comparison of the correlations between these (jointly fit) bias parameters and individual choice repetition probability. As in our previous analysis (Figure 4), individual choice repetition probability was more strongly predicted by drift than starting point bias (Figure 6c1).

A previous study of reward effects on speeded decisions reported that reward asymmetries induced supra-linear bias dynamics [263]. Temporal integration of a constant drift bias produces a linearly growing effective bias in the decision variable (Figure 1b), whereas integration of a ramping drift bias produces a supra-linear growth of effective bias (Figure 7a, left). In our data, a standard DDM with constant drift bias provided a better fit than DDMs with either a ramping drift bias, or a combination of constant and

ramping drift bias (Figure 7b2). Furthermore, in the latter (hybrid) model, the constant drift bias predicted individual choice repetition behavior better (Figure 7c2), in line with the constant accumulation bias inferred from the standard DDM fits. For the fits shown in Figure 7b2/c2, we used the same protocol as for the standard DDM, in which the time-varying sensory evidence fluctuations during stimulus presentation were replaced by their average over time to compute a single-trial drift rate (called 'default protocol', Methods section Extended bounded accumulation models: General assumptions and procedures). The same qualitative pattern of results also held for another fitting protocol ('dynamic protocol', see Methods), in which the time-varying sensory evidence was fed into the integrator (Δ AIC relative to no-history model: -1103, -985, -995, for constant drift bias, ramping drift bias, and hybrid, respectively; correlation with P(repeat): $r(30) = 0.5458$, $p = 0.0012$; $r(30) = 0.3600$, $p = 0.0429$ for constant and ramping drift bias, respectively). We next used this dynamic protocol for a set of more complex dynamical models.

It has been proposed that decision bounds might collapse over time, implementing an 'urgency signal' (Figure 6a, middle; [269, 270]). Indeed, adding collapsing bounds substantially improved our model fits (Figure 7b3). This indicates the presence of a strong urgency signal in this task, which had a relatively short stimulus presentation (750 ms) and a tight response deadline (1.25 s after stimulus offset). Critically, a history-dependent drift bias best fit the data (Figure 7b3) and captured individual choice repetition behavior (Figure 7c3) also in the DDM with collapsing bounds. In other words, while there is evidence for collapsing bounds in this dataset, our conclusion about the impact of history bias on decision formation does not depend on its inclusion in the model.

In the brain, a neural representation of the momentary sensory evidence feeds into a set of accumulators. These consist of circuits of excitatory and inhibitory populations of cortical neurons, which give rise to persistent activity and competitive winner-take-all dynamics [53, 271]. Under certain parameter regimes, these circuit dynamics can be reduced to lower-dimensional models [18, 144]. In such models, the effective accumulation time constant $1/\lambda$ (with λ being effective leak) results from the balance of leak within each accumulator (due to self-excitation and passive decay) and mutual inhibition between two accumulators encoding different choices [53]. Evidence accumulation can then be biased through an internal representation of the sensory input or through the way this sensory representation is accumulated (Figure 7a, right). We here used a reduced competing accumulator model, where the decision variable was computed as the difference of two leaky accumulators [68, 267, 272] to compare these two accumulation biases and a biased accumulator starting point.

We fit a family of bounded, leaky accumulator models, in which the starting point of the accumulators, their input, or their effective leak λ could be biased as a function of previous choice (Figure 7a, right). Note that a bias of the accumulator starting point would also translate into an accumulation bias, due to the model dynamics (see Methods section *Extended bounded accumulation models: General assumptions and procedures*). Even so, comparing this regime with other two biasing mechanism was informative. Also note that we here use the term 'leaky accumulator model' to denote that the model dynamics consisted of a free effective leak parameter l, without implying that $\lambda < 0$ (corresponding to activation decay). Our

fits allowed λ to take either negative ('forgetful' regime) or positive ('unstable' regime) values (Figure 7 – Figure Supplement 1d; see also [68]). Critically, in order to test for choice history-dependent accumulation bias, we allowed λ of each accumulator to vary as a function of the previous choice, before computing the difference between the two accumulator activations. Choice-history dependent biases in accumulator starting point or accumulator input were directly applied to the accumulator difference (akin to starting point and drift bias within the DDM). Due to the simplicity of its dynamics, the DDM cannot distinguish between input and leak bias. Indeed, when simulating behavior of leaky accumulator models with either of these two accumulation biases and fitting it with the DDM, both input and λ_{bias} loaded onto DDM drift bias (Figure 7 – Figure Supplement 3). Critically, the leaky accumulator with biased accumulator input best explained the data, among all the models considered (Figure 7b4). Furthermore, the individually estimated input bias predicted individual choice repetition (Figure 7c4). This suggests that choice history might specifically bias the internal representation of the sensory evidence feeding into the evidence accumulation process.

5.2.5 Dynamics of effective bias signal approximates rational combination of prior information current evidence

Taken together, fits and simulations of more complex models provide additional insight into the mechanism underlying choice history bias. They also corroborated the conclusion that choice history biases are mediated by a biased accumulation of evidence, rather than a biased starting point. As a final step, we estimated the time course of the effective bias, computed as the fraction of cumulative bias signal and bound height [273]. We simulated this signal based on the group average parameters for the best-fitting DDM and leaky accumulator models (Figure 7d). In the DDM and leaky accumulator (both with collapsing bound), the effective bias accelerated (Figure 7d).

The reader may notice that these (supra-linear) effective bias dynamics are similar to those predicted by the DDM with a ramping drift bias (Figure 7a, left). Thus, the observation that the latter model lost by a wide margin against the two models with more complex dynamics (Figure 7b, see also Materials and Methods) is likely due to features of the data other than the (relatively small) selective history bias. Specifically, the RT distributions were strongly shaped by the urgency signal incorporated by the bound collapse. In the overall best-fitting model (leaky accumulator with collapsing bounds and input bias, Figure 7b5), this effective bias depends on the combined effect of two non-linear signals: (i) the cumulative bias resulting from the accumulation of biased input and (ii) the hyperbolically collapsing bound. In the current fits, the effective bias was dominated by the strong bound collapse, but in different circumstances (with weaker urgency signal and for $\lambda < 0$), a biased input leaky accumulator can produce a decelerating effective bias. Combination of a biased input with some starting point and or leak bias can further change the dynamics. The key observation is that, regardless of the modeling framework used, we identified an effective bias signal that grew steadily throughout decision formation, in line with the main conclusion drawn from the basic fits of the standard DDM.

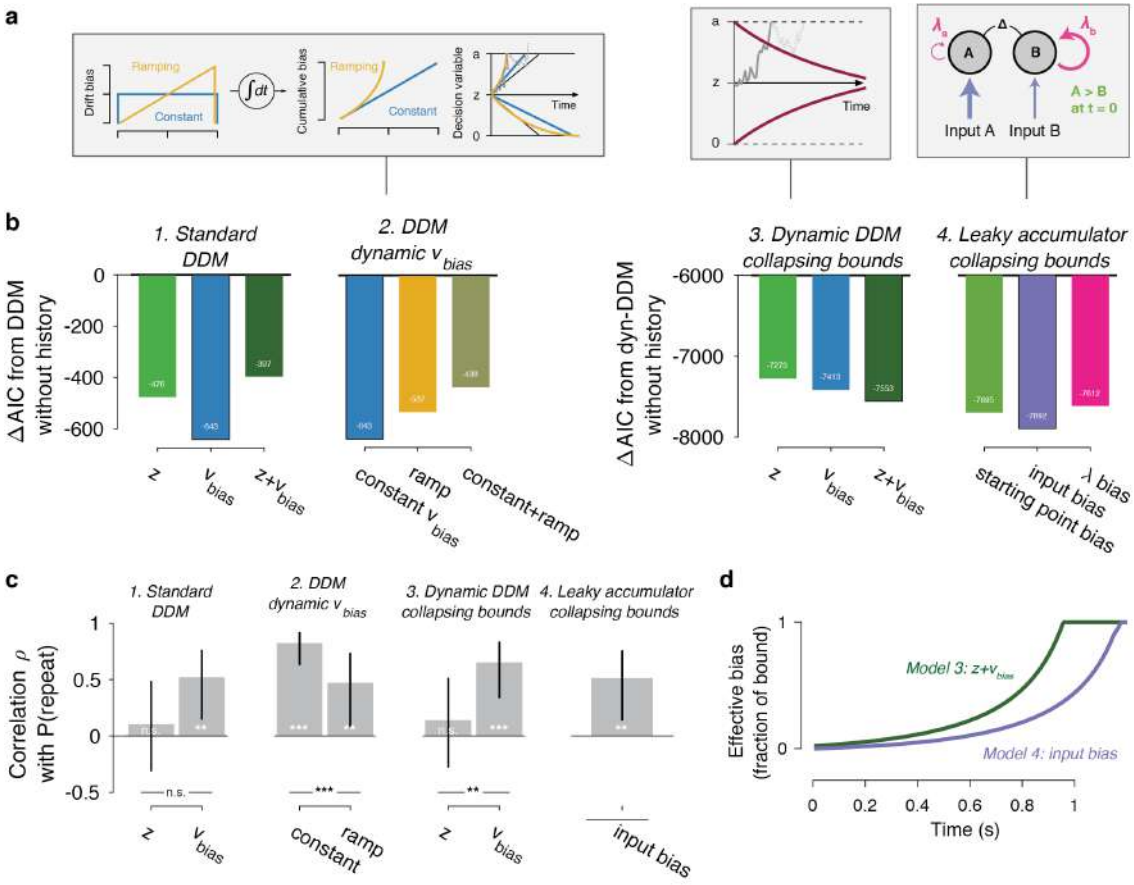


Figure 7: Extended dynamic models of biased evidence accumulation. (a) Model schematics. See main text for details. In the third panel from the left, the stimulus-dependent mean drift is shown in black, overlaid by the biased mean drift in color (as in Figure 1a,b). **(b)** AIC values for each history-dependent model, as compared to a standard (left) or dynamic (right) DDM without history. The winning model (lowest AIC value) within each model class is shown with a black outline. In 3-4, the dashed line indicates the AIC value of the standard DDM without history that is the baseline in 1-2. **(c)** Correlation (Spearman's ρ) of parameter estimates with individual repetition behavior, as in Fig. 4b. Error bars, 95% confidence interval. *** $p < 0.0001$, ** $p < 0.01$, n.s. $p > 0.05$. **(d)** Within-trial time courses of effective bias (cumulative bias as a fraction of the decision bound) for the winning DDM and leaky accumulator models. Effective bias time courses are indistinguishable between both dynamical regimes $\lambda < 0$ and $\lambda > 0$) and are averaged here.

Our results are in line with the idea the impact of choice history bias on decision formation grows as a function of elapsed time. Because prior information (here: about theprevious choice) does not change over time, this observation might be surprising. Yet, previous work has identified a principled rationale for such a time-dependent combination of prior and evidence. When evidence reliability changes from trial to trial, prior information (bias) should be weighted more strongly when sensory evidence is unreliable [273, 274]. This can be achieved by increasing the weight of the prior throughout the trial, using elapsed time as a proxy for evidence reliability. This prediction was confirmed experimentally for explicit manipulations of prior probability of the choice options [273]. Indeed, within the framework of the DDM, this way of combining prior information with current evidence maximizes reward rate [274, 275]. Only when evidence reliability is constant across trials should prior information be incorporated as a static bias (i.e., starting point). Evidence reliability likely varied from trial to trial across all our experiments [274],

due to variations in the external input (i.e., mean drift rate in the DDM), originating from stochastically generated stimuli, or internal factors (i.e., drift rate variability in the DDM), such as the inherent variability of sensory cortical responses [69, 276]. In particular, the dataset from Figure 7 entailed strong trial-to-trial variations in the external input (Figure 7 – Figure Supplement 1). Thus, the dynamics of the effective bias signal uncovered in Figure 7d suggest that participants combined prior information with current evidence in a rational fashion.

5.3 Discussion

Quantitative treatments of perceptual decision-making commonly attribute trial-to-trial variability of overt choice to noise in the decision computation [3, 5, 277]. Those accounts typically assume that systematic decision biases remain constant over time. Instead, the choice history biases studied here vary continuously over the course of the experiment, as a function of the previous choices (and choice outcome information). Our current results indicate that choice history explains trial-to-trial variability specifically in evidence accumulation, in a number of widely used perceptual choice tasks. Ignoring such trial-to-trial variations will lead to an overestimation of the noise in the evidence accumulation process and resulting behavior.

History biases in perceptual choice have long been known in perceptual psychophysics [14] and neuroscience [229]. However, the underlying dynamical mechanism has remained elusive. We here show that individual differences in overt choice repetition behavior are explained by individual differences in the degree to which choices bias the evidence accumulation, not the starting point, of subsequent decisions. This accumulation bias is associated with choices made several trials into the past, and it grows steadily as the current decision unfolds. This insight calls for a revision of current models of choice history biases [239, 240].

It is instructive to relate our results with previous studies manipulating the probability of the occurrence a particular category (i.e., independently of the sequence of categories) or the asymmetry between rewards for both choices. Most of these studies explained the resulting behavioral biases in terms of starting point shifts [244, 249, 260, 261, 278], but only for decisions without time pressure [263]. Yet, one study with variations of evidence strength found an effect of asymmetric target probability on accumulation bias [273] similar to the one we here identified for choice history. In all this previous work, biases were under experimental control: Probability or reward manipulations were signaled via explicit task instructions or single-trial cues (in humans) or block structure (in animals). By contrast, the choice history biases we studied here emerge spontaneously and in an idiosyncratic fashion (Figure 2e), necessitating our focus on individual differences.

Our modeling addressed the question of how prior information is combined with new evidence during decision formation (see in particular the section *Dynamics of effective bias signal approximates rational combination of prior information with current evidence*). But why did participants use choice history as a prior for their decisions? In all our experiments, the sensory evidence was uncorrelated across trials – as

is the case in the majority of perceptual choice tasks used in the literature. Thus, any history bias can only reduce performance below the level that could be achieved, given the observer's sensitivity. It may seem irrational that people use history biases in such settings. However, real-world sensory evidence is typically stable (i.e., auto-correlated) across various timescales [239]. Thus, people might (erroneously) apply an internal model of this environmental stability to randomized laboratory experiments [239], which will push them towards choice repetition or alternation [241]. Indeed, people flexibly adjust their choice history biases to environments with different levels of stability [208, 241, 279], revealing the importance of such internal models on perceptual decision-making. In sum, with our conclusions from the time course of the effective bias signal, these considerations suggest that participants may have applied a rational strategy, but based on erroneous assumptions about the structure of the environment.

While we found that choice history-dependent variations of accumulation bias were generally more predictive of individual choice repetition behavior, the DDM starting point was consistently shifted away from the previous response for a majority of participants (i.e., negative values along x-axis of Figures 4a). This shift was statistically significant in 3 out or 6 datasets (Figure 4 – Figure Supplement 5a), and might explain the advantage of the dual parameter model over the pure drift-bias model in our model comparisons (Figure 3a). The starting point shift may be due to at least two scenarios, which are not mutually exclusive. First, it might reflect a stereotypical response alternation tendency originating from dynamics in motor cortex – for example, a post-movement "rebound" of beta-band oscillations [280]. Indeed, previous work found that beta rebound is related response alternation in a perceptual choice task, which precluded (in contrast to our tasks) motor planning during evidence accumulation [233]. This stereotypical response alternation tendency (via starting point) may have conspired with the more flexible history bias of evidence accumulation (via drift) to shape overall choice behavior. Because starting point shifts will predominantly contribute to fast decisions, this scenario is consistent with the average choice alternation tendency we observed for RTs < 600ms (Figure 4 – Figure Supplement 5c). Because the response alternation tendency in motor cortex is likely to be induced only by the immediately preceding response, this scenario is also consistent with the shorter timescales we estimated for the starting point effects (1.39 trials) than the drift rate effects (2.38 trials; Figure 6a, black exponential fit lines). Second, the starting point shift may also reflect decision dynamics more complex than described by the standard DDM: non-linear drift biases (Figure 7 – Figure Supplement 2, third column) or biases in the leak of decision accumulators (Figure 7 – Figure Supplement 3, third column). Both give rise to opposite effects on drift bias and starting point bias when fit with the standard DDM, thus yielding negative correlations between DDM starting point and drift bias estimates. Such negative correlations were present in our data, but weak and not statistically significant (Spearman's rho -0.4130 to -0.0757, combined BF_{10} 0.0473). It is possible that both of the scenarios discussed here conspired to yield the starting point effects observed in model comparisons and individual parameter estimates. Future work is needed to illuminate this issue, for example through manipulating decision speed and/or the delays between subsequent motor responses, and modeling choice-related neural dynamics in the motor cortex.

We propose that choice history biases evidence accumulation, but there are alternative scenarios. First, it is possible that participants' choices were due to computations altogether different from those incorporated in the bounded accumulation models assessed here. All our models imply simple neural accumulators with persistent activity. At least on a subset of trials, participants may make fast guesses [281], or engage in automatic decision processing [282, 283] or post-accumulation biases [284]. Also, the decision computation may entail noise-driven attractor dynamics [271, 285] possibly with sudden 'jumps' between neural activity states [286], instead of linear accumulation to a threshold level. Even if the accumulation dynamics postulated in our models cannot be reduced to the dynamics of single neurons, the history-dependent accumulation bias we inferred here would constitute a valid description of the collective computational properties of the neural system producing choice behavior. Second, within bounded accumulation models, any directed change in the decision variable can be mimicked by some selective (i.e. asymmetric) change in one of the decision bounds. For example, in simulations we find that combining the DDM with a linearly collapsing bound for the favored choice and a linearly expanding bound for the other choice has the same effect on choice fractions and RT distributions as a drift bias (data not shown). We are not aware of any empirical evidence for such asymmetric changes in decision bounds. Decision-related cortical ramping activity seems to always reach a fixed level just before motor response, irrespective of prior probabilities [273] or speed-accuracy trade-offs [42, 273]. Instead, the build-up of this activity is biased by prior information [273].

A plausible mechanism underlying the choice history-dependent shift in accumulation bias is a bias of the neural representations of the sensory evidence towards (or away from) a previously selected category [210, 235, 287]. This is precisely the 'input bias' scenario entailed in our best fitting model (Figure 7). The primate brain is equipped with powerful machinery to bias sensory representations in a top-down fashion [288, 289]. In the laboratory, these top-down mechanisms have been probed by explicitly instructing subjects to shift their attention to a particular sensory feature or location. Such instructions induce biased activity states in regions of prefrontal and parietal association cortex, which are propagated down the cortical hierarchy to sensory cortex via selective feedback projections, where they boost the corresponding feature representations and suppress other feature representations [288]. The same prefrontal and parietal regions accumulate sensory evidence towards choices and seem to carry choice history signals. It is tempting to speculate that choice history signals in these regions cause the same top-down modulation of sensory cortex as during explicit manipulations of attention. In other words, agents' choices might be one factor directing their top-down attention under natural conditions, in a way analogous to explicit attention cues in laboratory tasks. An alternative, but related possibility is that the direction of selective attention fluctuates spontaneously during the course of a perceptual choice experiment, preferentially sampling features supporting one choice for a streak of trials, and then switching to sampling support for the other category. The corresponding top-down modulations would bias evidence accumulation and choice in a serially correlated fashion. These ideas are not mutually exclusive and can be tested by means of multi-area neurophysiological recordings combined with local perturbations.

A growing body of evidence points to the interplay of multiple timescales for neural computation in the primate cortex. One line of behavioral work has revealed effective (within-trial) evidence accumulation over timescales ranging from a few hundred milliseconds [252, 253] to several seconds [43, 132, 290]. Another line, including the current study, revealed the slow accumulation of internal decision variables or external outcome information across trials (tens of seconds) to build up time-varying biases, or priors [13, 208, 232, 291]. Relatedly, neurophysiological work on ongoing activity has inferred multiple, hierarchically organized, timescales in different cortical regions [238, 292–295]. The history-dependent evidence accumulation biases we have uncovered here might index the interplay between these different effective timescales with long-timescale accumulators at higher stages biasing short-timescale accumulators at intermediate stages of the cortical hierarchy.

5.4 Materials and Methods

5.4.1 Datasets: behavioral tasks and participants

We analyzed six different datasets, four of which were previously published. These spanned different modalities (visual or auditory), decision-relevant sensory features (motion direction, contrast, tone presence, motion coherence), and tasks (detection or discrimination). In each dataset, the number of participants was determined to allow for robust estimation of the original effects of interest. No participants were excluded from the analyses.

Those tasks where the decision-relevant sensory evidence was presented until the observer generated a response were called response time (RT) tasks; those tasks where the sensory evidence is presented for a fixed duration, and its offset cues the observer's response, were called fixed duration (FD) tasks in line with the terminology from [296]. These two protocols have also been termed 'free response protocol' and 'interrogation protocol' [18]. In all datasets, stimulus strength (i.e., decision difficulty) was kept constant, or varied systematically across levels, within all main experimental sessions that were used for fitting the DDM.

2AFC visual motion discrimination task (RT) These data were previously published [34], and are available at `https://doi.org/10.5061/dryad.tb542`. The study was approved by the ethics committee of the Leiden University Cognitive Psychology department, and all subjects provided written informed consent prior to taking part. Twenty-six observers (22 women and 4 men, aged 18-29) performed a motion direction (left vs. right) discrimination task. Stationary white dots were presented on a black screen for an interval of 4.3-5.8 s. After this fixation interval, the decision-relevant sensory evidence was presented: some percentage of dots (the 'motion coherence' level) moved to the left or the right. The coherence was individually titrated to yield an accuracy level of 85% correct, estimated from a psychometric function fit, before the start of the main experiment, and kept constant afterwards. The moving dots were presented until observers indicated their choice with a button press. After the response, the fixation

cross changed color for 700 ms to indicate single-trial feedback. Each observer performed 500 trials of the task (one session). We refer to this task as 'Visual motion 2AFC (RT)'.

2AFC VISUAL MOTION DISCRIMINATION TASK (FD) Thirty-two participants (aged 19-35 years, 43 women and 21 men) participated in the study after giving their informed consent. The experiment was approved by the ethical review board of the University Medical Center Hamburg-Eppendorf (PV4714). Observers performed a fixed duration version of the random dot motion discrimination (up vs. down) task in the MEG scanner. White dots were displayed on a grey background screen, with a density of 6 dots/degree2, resulting in 118 dots on the screen at each frame. The stimuli were confined to a circle of 2.5° radius, which was placed in the lower half of the visual field at 3.5° from the fixation. After a fixation interval of 0.75–1.5s, random dot motion stimuli (0, 3, 9, 27 or 81% motion coherence) were displayed for 750 ms. Signal dots moved with a speeds of 11.5 degree/s, and noise dots were randomly displaced within the circle on each frame. We used the single-trial dot coordinates to construct time courses of fluctuating external evidence (see Methods section *Motion energy filtering and psychophysical kernels*; Figure 7 – Figure Supplement 1a-c). Observers received auditory feedback 1.5-2.5s after their response, and the ISI started 2-2.5s after feedback. Observed performed 1782 trials over 3 sessions, in which the stimulus transition probability varied (0.2, 0.5 or 0.8) between blocks of 99 trials. To maximize trial counts for the non-hierarchical leaky accumulator fits, we here collapsed across blocks. We refer to this task as 'Visual motion 2AFC (FD)'.

2AFC VISUAL MOTION DISCRIMINATION 2IFC TASK (FD): DATA SET 1 These data were previously published in [125], and are available at http://dx.doi.org/10.6084/m9.figshare.4300043. The ethics committee at the University of Amsterdam approved the study, and all observers gave their informed consent before participation. Twenty-seven observers (17 women and 10 men, aged 18-43) performed a two-interval motion coherence discrimination task. They viewed two consecutive intervals of random dot motion, containing coherent motion signals in a constant direction towards one of the four diagonals (counterbalanced across participants) and judged whether the second test interval (variable coherence) contained stronger or weaker motion than the first reference (constant coherence) interval. After a fixation interval of 0.5-1s, they viewed two consecutive intervals of 500 ms each, separated by a delay of 300-700 ms. The decision-relevant sensory evidence (i.e., the difference in motion coherence between intervals), was chosen pseudo-randomly for each trial from the set (0.625, 1.25, 2.5, 5, 10, 20, 30%). Observers received auditory feedback on their choice after a delay of 1.5-2.5s. After continuing to view noise dots for 2-2.5 s, stationary dots indicated an inter-trial interval. Observers self-initiated the start of the next trial (range of median inter-trial intervals across observers: 0.68–2.05 s). Each observer performed 2500 trials of the task, divided over five sessions. We refer to this task as 'Visual motion 2IFC (FD) #1'.

2AFC VISUAL MOTION DISCRIMINATION 2IFC TASK (FD): DATA SET 2 Sixty-two participants (aged 19-35 years, 43 women and 19 men) participated in the study after screening for psychiatric, neuro-

logical or medical conditions. All subjects had normal or corrected to normal vision, were non-smokers, and gave their informed consent before the start of the study. The experiment was approved by the ethical review board of the University Medical Center Hamburg-Eppendorf (PV4648). Observers performed 5 sessions, of which the first and the last took place in the MEG scanner (600 trials divided over 10 blocks per session) and the three sessions in between took place in a behavioral lab (1500 trials divided over 15 blocks per session). The task was as described above for 'Visual motion 2IFC (FD) #1', with the following exceptions. The strength of the decision-relevant sensory evidence was individually titrated to an accuracy level of 70% correct, estimated from a psychometric function fit, before the start of the main experiment and kept constant for each individual throughout the main experiment. Each stimulus was presented for 750 ms. In the MEG sessions, auditory feedback was presented 1.5-3 s after response, and an inter-trial interval with stationary dots started 2-3 s after feedback. Participants initiated the next trial with a button press (across-subject range of median inter-trial interval duration: 0.64 to 2.52 s, group average: 1.18 s). In the training sessions, auditory feedback was presented immediately after the response. This was followed by an inter-trial interval of 1 s, after which the next trial started. In this experiment, three sub-groups of observers received different pharmacological treatments prior to each session, receiving placebo, atomoxetine (a noradrenaline reuptake inhibitor), or donepezil (an acetylcholinesterase inhibitor). These groups did not differ in their choice history bias and were pooled for the purpose of the present study (Figure 4 – Figure Supplement 2). We refer to this task as 'Visual motion 2IFC (FD) #2'.

Visual contrast yes/no detection task (RT) These data were previously published [40], and are available at https://doi.org/10.6084/m9.figshare.4806559. The ethics committee of the Psychology Department of the University of Amsterdam approved the study. All participants took part after giving their written informed consent. Twenty-nine observers (14 women and 15 men, aged 18–38) performed a yes/no contrast detection task. During a fixation interval of 4-6 seconds, observers viewed dynamic noise (a binary noise pattern that was refreshed each frame, at 100 Hz). A beep indicated the start of the decision-relevant sensory evidence. On half the trials, a vertical grating was superimposed onto the dynamic noise; on the other half of trials, only the dynamic noise was shown. The sensory evidence (signal+noise or noise-only) was presented until the observers reported their choice (yes, grating was present; or no, grating was absent), or after a maximum of 2.5s. The signal contrast was individually titrated to yield an accuracy level of 75% correct using a method of constant stimuli before the main experiment, and kept constant throughout the main experiment. Observers performed between 480–800 trials over 6-10 sessions. Six observers in the original paper [40] performed a longer version of the task in which they also reported their confidence levels and received feedback; these were left out of the current analysis, leaving twenty-three subjects to be included. We refer to this task as 'Visual contrast yes/no (RT)'.

Auditory tone yes/no detection task (RT) These data were previously published [33], and are available at https://doi.org/10.6084/m9.figshare.4806562. All subjects gave written in-

formed consent. The ethics committee of the Psychology Department of the University of Amsterdam approved the experiment. Twenty-four observers (20 women and 4 men, aged 19–23) performed an auditory tone detection task. After an inter-trial interval of 3-4 seconds, decision-relevant sensory evidence was presented: on half the trials, a sine wave (2 KHz) superimposed onto dynamic noise (so-called TORCS; [76]) was presented; on the other half of trials only the dynamic noise was presented. The sensory evidence was presented until the participant reported their choice button press or after a maximum of 2.5s. No feedback was provided. Each individual's signal volume was titrated to an accuracy level of 75% correct using an adaptive staircase procedure before the start of the main experiment, and kept constant throughout the main experiment. Participants performed between 1320 and 1560 trials each, divided over two sessions. We refer to this task as 'Auditory yes/no (RT)'.

5.4.2 Model free analysis of sensitivity and choice history bias

We quantified perceptual sensitivity in terms of signal detection-theoretic d' [24]:

$$d' = \Phi^{-1}(H) - \Phi^{-1}(FA) \tag{1}$$

where Φ was the normal cumulative distribution function, H was the fraction of hits and FA the fraction of false alarms. In the 2AFC and 2IFC datasets, one of the two stimulus categories was arbitrarily treated as signal absent. Both H and FA were bounded between 0.001 and 0.999 to allow for computation of d' in case of near-perfect performance [297]. We estimated d' separately for each individual and, for the two datasets with varying difficulty levels, for each level of sensory evidence.

We quantified individual choice history bias in terms of the probability of repeating a choice, termed P(repeat), regardless of the category of the (previous or current) stimulus. This yielded a measure of bias that ranged between 0 (maximum alternation bias) and 1 (maximum repetition bias), whereby 0.5 indicated no bias.

5.4.3 Conditional bias functions

For each variant of the model and each dataset, we simulated data using the best-fitting parameters. Specifically, we simulated 100 responses (choices and RTs) for each trial performed by the observers. These predicted patterns for the 'baseline model' (without history-dependent bias parameters) were first used to compare the observed and predicted patterns of choices and RTs (Figure 3 – Figure Supplement 2).

We also used the simulated data, as well as the participants' choices and RTs, to visualize specific features in our data that distinguish the different biased models [255]. Specifically, we computed conditional bias functions [249] that visualize choice history bias as a function of RTs. Each choice was recoded into a repetition (1) or alternation (0) of the previous choice. We then expressed each choice as being either in line with or against the observer's individual bias (classified into 'repeaters' and 'alternators' depending on choice repetition probability). Note that given the transformation of the data (sign-flipping the bias data

for alternators in order to merge the two groups), the fact the average P(bias) > 0.5 is trivial, and would occur for any generative model of history bias. Conditional bias functions instead focus on the effect of choice history bias as a function of time within each trial, the shape of which distinguishes between different bias sources (Figure 1c).

To generate these conditional bias functions, we divided each (simulated or real) observer's RT distribution into five quantiles (0.1, 0.3, 0.5, 0.7 and 0.9) and computed the fraction of biased choices within each quantile. The shape of the conditional bias functions for models with z and vbias confirm that z predominantly produces biased choices with short RTs, whereas vbias leads to biased choices across the entire range of RTs (Figure 3b).

5.4.4 Drift diffusion model (DDM) fits

General This section describes the general DDM, with a focus on the biasing mechanisms described in Results and illustrated in Figure 1 [25]. Ignoring non-decision time, drift rate variability, and starting point variability (see below), the DDM describes the accumulation of noisy sensory evidence as follows:

$$dy = s \cdot v \cdot dt + cdW \tag{2}$$

where y is the decision variable (gray example traces in Figure 1), s is the stimulus category (coded as $[-1, 1]$), v is the drift rate, and cdW is Gaussian distributed white noise with mean 0 and variance c^2dt [18]. In an unbiased case, the starting point of the decision variably $y(0) = z$, is situated midway between the two decision bounds 0 and a:

$$y(0) = z = \frac{a}{2} \tag{3}$$

where a is the separation between the two decision bounds. A bias in the starting point is implemented by an additive offset z_{bias} from the midpoint between the two bounds (Figure 1a):

$$y(0) = z = \frac{a}{2} + z_{bias} \tag{4}$$

A drift bias can be implemented by adding a stimulus-independent constant v_{bias}, also referred to as drift bias [25], to the (stimulus-dependent) mean drift (Figure 1b). This adds a bias to the drift that linearly grows with time:

$$dy = (s \cdot v + v_{bias})dt + cdW \tag{5}$$

We allowed both bias parameters to vary as a function of observers' previous choice, to test their relative contributions to the individual differences in overt choice history biases. These two biasing mechanisms result in the same (asymmetric) fraction of choices, but they differ in terms of the resulting shapes of RT

distributions (Figure 1). In previous work, z_{bias} and v_{bias} have also been referred to as 'prior' and 'dynamic' bias [105] or 'judgmental' and 'perceptual' bias [262].

Estimating HDDM bias parameters We used hierarchical drift diffusion modeling as implemented in the HDDM toolbox [77] to fit the model and estimate its parameters. As recommended by the HDDM toolbox, we specified 5% of responses to be contaminants, meaning they arise from a process other than the accumulation of evidence – for example, a lapse in attention [298]. We fit the DDM to RT distributions for the two choice categories, conditioned on the stimulus category for each trial (s in eq. 2) – a procedure referred to as 'stimulus coding'. This fitting method deviates from a widely used expression of the model, where RT distributions for correct and incorrect choices are fit (also called 'accuracy coding'). Only the former can fit decision biases towards one choice over the other.

First, we estimated a model without history-dependent bias parameters. Overall drift rate, boundary separation, non-decision time, starting point, and drift bias were estimated for each individual (Figure 3 – Figure Supplement 1). Across-trial variability in drift rate and starting point were estimated at the group-level only [112]. For the datasets including variations of sensory evidence strength (Visual motion 2AFC (FD) and Visual motion 2IFC (FD) #1), we separately estimated drift rate for each level of evidence strength. This model was used to confirm that the DDM was able to fit all datasets well, and to serve as a baseline for model comparison.

Second, we estimated three different models of history bias, allowing (i) starting point, (ii) drift or (iii) both to vary as a function of the observer's immediately preceding choice (thus capturing only so-called first-order sequential effects; [243, 246]). The effect of the preceding choice on each bias parameter was then termed its 'history shift'. For example, for the visual motion direction discrimination task we separately estimated the starting point parameter for trials following 'left' and 'right' choices. The difference between these two parameters then reflected individual observers' history shift in starting point, computed such that a positive value reflected a tendency towards repetition and a negative value a tendency towards alternation. The history shift in drift bias was computed in the same way.

HDDM regression models We estimated the effect of up to 6 previous stimuli and choices on history bias using a HDDM regression model. We first created a design matrix X with dimensions $trials \times (2 * lags)$, which included pairs of regressors coding for previous stimuli and choices (coded as −1,1), until (and including) each model's lag. Two distinct replicas of X were then used as design matrices to predict drift bias (X_v) and starting point (X_z). Drift bias was defined as $v \sim 1 + s + X_v$, where 1 captured an overall bias for one choice over the other and s indicated the signed stimulus strength. Starting point was defined as $z \sim 1 + X_z$, with a logistic link function $\frac{1}{1+e^{-X}}$.

After fitting, parameter estimates were recombined to reflect the effect of previous correct (choice + stimuli) or error (choice − stimuli) trials. We sign-flipped the weight values for alternators (i.e. those participants with a repetition tendency at lag one < 0.5); this makes all the panels in Figure 6b interpretable as a change in each parameter in the direction of individual history bias.

HDDM fitting procedures The HDDM [77] uses Markov-chain Monte Carlo sampling for generating posterior distributions over model parameters. Two features of this method deviate from more standard model optimization. First, the 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. Second, 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 (Figure 3 – Figure Supplement 1); [77]). In practice, this results in more stable parameter estimates for individual subjects, who are constrained by the group-level inference. Note that we also repeated our model fits with more traditional G2 optimization [298] and obtained qualitatively identical results (Figure 4 – Figure Supplement 1a).

For each variant of the model, we ran 30 separate Markov chains with 5000 samples each. Of those, half were discarded as burn-in and every second sample was discarded for thinning, reducing autocorrelation in the chains. This left 1250 samples per chain, which were concatenated across chains. Individual parameter estimates were then estimated from the posterior distributions across the resulting 37500 samples. All group-level chains were visually inspected to ensure convergence. Additionally, we computed the Gelman-Rubin $\hat{R}$ statistic (which compares within-chain and between-chain variance) and checked that all group-level parameters had an $\hat{R}$ between 0.98–1.05.

Formal comparison between the different model variants was performed using the Akaike Information Criterion [254]: $AIC = -2L + 2k$, where L is the total log likelihood of the model and k denotes the number of free parameters. The AIC was computed for each observer, and summed across them. Lower AIC values indicate a better fit, while taking into account the complexity of each model. A difference in AIC values of more than 10 is considered evidence for the winning model to capture the data significantly better. The conclusions drawn from AIC hold also when using the Deviance Information Criterion for the hierarchical models.

5.4.5 Motion energy filtering and psychophysical kernels

For the Visual motion 2AFC (FD) dataset, we used motion energy filtering (using the filters described in [210]) to reconstruct the time-course of fluctuating sensory evidence over the course of each individual trial, averaging over the spatial dimensions of the display (Figure 7 – Figure Supplement 1a,b). These single-trial traces then served as the time-resolved input to a set of extended DDM and leaky accumulator models (Figure 7). Specifically, filtering the stimuli at 60 Hz (the refresh rate of the LCD projector) resulted in 45 discrete samples for the 750 ms viewing period of each trial. The first 13 samples of the motion energy filter output (first 200 ms of the viewing interval) corresponded to the 'rise time' of the filter [252], yielding outputs that were a poor representation of the actual motion energy levels (see also Figure 7 – Figure Supplement 1a). In order to prevent those uninterpretable filter outputs from contributing, we discarded the first 15 samples (250 ms) before model fitting (see below). Using constant interpolation, we expanded the remaining 30 samples onto 150 samples, which, given that the simulation Euler step was 5

ms (dt = 0.005), corresponded to a 750 ms long input time series. In the model descriptions below we denote the input time series with $M = \{M_t : t \in T\}$ and $T = \{1, 2, ..., 150\}$.

We also used these motion energy traces to construct so-called psychophysical kernels. Within each stimulus identity (motion direction and coherence, excluding the easiest 81% coherence trials), we subtracted the average motion energy traces corresponding to 'up' vs. 'down' choices. The resulting trace represents the excess motion energy that drives choices, over and above the generative stimulus coherence (Figure 7 – Figure Supplement 1c).

5.4.6 Extended bounded accumulation models

General assumptions and procedures

In the 2AFC (FD) visual motion experiment participants viewed the stimulus for 0.75 s (hereafter called 'viewing period') and could respond only after the stimulus offset. This required specifying the input to the evidence accumulation process. In the models described below we used separate simulation protocols, based on different assumptions about this input. In the 'dynamic' protocol, where the input was the time-varying sensory evidence from each trial, the accumulation process was assumed to start at stimulus onset, and responses could happen during the motion viewing interval. The average activity of the accumulator(s) at stimulus offset served as input for accumulation during the post-offset period. For fitting models using this protocol, empirical RTs were calculated relative to the stimulus onset. Motion energy estimates were used as time-resolved input to the model.

By contrast, in the 'default' protocol, the motion energy fluctuations were averaged across the viewing interval excluding the filter rise time (i.e., from 250 to 750 s after stimulus offset), and the average motion energy was then used as a single-trial drift rate for the accumulation process. In other words, the accumulation-to-bound dynamics only took place during the post-offset period. Accordingly, when fitting models with this protocol, the empirical RTs were calculated relative to stimulus offset. Using this protocol was necessary for replicating our basic result from the standard DDM fits: For the 'dynamic' protocol, any starting point bias would turn into a drift bias because it would feed into accumulation process after stimulus offset, precluding the comparison between the two forms of bias. Thus, we used only the default protocol for the standard DDM fits, which aimed at differentiating between starting point and accumulation biases. For comparison, we also used the same simulation protocol when fitting an extended DDM with a both a constant and a ramping component in the drift bias (see below). We then switched to the more realistic dynamic protocol for the subsequent models with more complex dynamics.

The AIC scores of models using the default protocol were generally lower (better) compared to the respective models that used the dynamic protocol. This difference is likely due to the fact that the dynamic protocol is more constrained by using as input to the models the exact motion energy traces rather than just their mean for each trial. AIC is blind to such latent flexibility differences that do not map onto differences in number of parameters. Thus, AIC may have "under-penalized" models in the default protocol relative to those in the dynamic protocol.

In all models and in both simulation protocols, model predictions were derived via Monte Carlo simulation. The variance of the processing noise was set to $c^2 = 1$. One simulation time-step corresponded to 5 ms (Euler step, dt = 0.005). Finally, in the standard protocol the accumulation process could last for a maximum of 300 time-steps (or 1500 ms) and in the dynamic protocol for a maximum of 450 time-steps (or 2250 ms). After these time points, the process timed-out and a response was assigned to the alternative according to the state of the diffusion variable (e.g., in the standard DDM right if y > a/2 and left if y < a/2).

DDM variants with default simulation protocol

For all basic DDM variants described in this section, we used the default simulation protocol: the time-averaged motion energy for each trial provided the drift-rate (v) driving the subsequent diffusion process. DDM models had 5 generic parameters: threshold (a), noise scaling (g), non-decision time (Ter), drift-rate variability (sv) and starting-point variability (sz).

Naïve DDM. We denote with y the state of the diffusion variable. At time 0:

$$y(0) = z = \frac{a}{2} + U(-sz, sz) \tag{6}$$

where U was a uniform random variable (rectangular distribution) in the $(-sz, sz)$ range. The evolution of y was described by:

$$dy = g \cdot \ddot{v} \cdot dt + cdW \tag{7}$$

Above, g was the scaling parameter that controls the signal-to-noise-ration (given that c is fixed at 1). The variable $\ddot{v}$ was the effective drift-rate, i.e. a Gaussian variable with $N(m, sz^2)$ where sz was the drift-rate variability and m was the average of the motion energy on each trial. A response was generated when the decision variable y exceeded a (right choice) or surpassed 0 (left choice). The moment that either of these boundaries was crossed plus a non-decision time Ter, determined the per-trial RT.

Starting point DDM. This model was the same as the naïve model but with an extra parameter z_{bias} such that at time 0:

$$y(0) = z = \frac{a}{2} + U(-sz, sz) + z_{bias} \cdot prev \tag{8}$$

The variable $prev$ here encoded the previous choice (1: right, -1: left). If $_{bias}$ was positive the model implemented repetition and if negative it implemented alternation.

Drift bias DDM. Same as the naïve model but with an extra biasing parameter v_{bias} such that:

$$dy = (g \cdot \ddot{v} + v_{bias} \cdot prev)dt + cdW \tag{9}$$

Hybrid DDM. This version combined the starting point DDM and drift bias DDM using two biasing parameters.

Simple Ramping DDM. This model was the same as the naïve model but with an extra parameter s_{ramp} such that:

$$dy = (g \cdot \ddot{v} + \frac{s_{ramp} \cdot t \cdot prev}{t_{max}})dt + cdW \tag{10}$$

where t denoted time elapsed in terms of Monte-Carlo time-steps and t_{max} = 300 time-steps, which was the maximum duration that a given trial could run for.

Hybrid Ramping DDM. Same as the naïve model but with 2 extra parameters s_{ramp} and $s_{constant}$ such that:

$$dy = (g \cdot \ddot{v} + (s_{constant} + \frac{s_{ramp} \cdot t}{t_{max}})prev)dt + cdW \tag{11}$$

This model thus implemented a drift bias that is nonzero at the start of the trial ($s_{constant}$), and also linearly increases until the end of the trial (with slope s_{ramp}).

Extended models with dynamic simulation protocol

For all subsequently described models, we used the dynamic simulation protocol (see section *General Assumptions and Procedures*), with the motion energy time courses serving as input to the accumulation process. To illustrate the details of the dynamic protocol, we next describe how the decision variable was updated in the case of the naïve DDM. The decision variable during the viewing period evolved according to the following differential equation:

$$dy = g \cdot M_t \cdot dt + cdW \tag{12}$$

where M_t was the value of the input signal at time t. Following stimulus offset (at $t = T$), after 150 time-steps, the diffusion variable carried on being updated as follows:

$$dy(t) = \frac{y(T)}{T} + cdW \tag{13}$$

In other words, after the stimulus disappeared, accumulation was driven by the average evidence accumulated up to the point of stimulus offset. This post-stimulus accumulation could continue for a maximum of 300 extra time-steps, at which point the process timed out.

Simple and Hybrid Ramping DDM. This model was the same as the above Simple and Hybrid Ramping DDMs, only now fit by using the dynamic simulation protocol (i.e. the ramping drift-criterion bias is applied for the viewing period only and, following stimulus offset, the decision variable is updated according to equation 13).

Dynamic DDM with Collapsing Bounds. In the "collapsing bounds" DDM models, a response was generated when the diffusion variable (y) exceeds b_{up} (right choice) or surpasses b_{down} (left choice). The two thresholds, b_{up} and b_{down}, vary across time as follows:

$$b_{up}(t) = |a - a\frac{t}{t+c}|_{a/2}^{a} \tag{14a}$$

$$b_{down}(t) = |a\frac{t}{t+c}|_{0}^{a/2} \tag{14b}$$

In the above, the notation $|x|_{min}^{max}$ indicates that x was clamped such that $x \in [min, max]$. The moment that either of these boundaries was reached, plus a non-decision time Ter, determined the per-trial RT. The dynamic DDM model had 5 basic parameters: threshold initial value (a), threshold collapse rate (c), noise scaling (g), non-decision time (Ter), and starting-point variability (sz).

Starting point dynamic DDM. Here, the state of the diffusion variable was initialized according to equation 8. Thus, the starting point model had 6 free parameters (the 5 basic ones plus the starting point bias, z_{bias}).

Drift-bias dynamic DDM. The diffusion variable at time 0 was initialized according to equation 8. Also, the diffusion variable in the viewing period was not updated according to equation 9 but according to:

$$dy = (g \cdot M_t + v_{bias} \cdot prev)dt + cdW \tag{15}$$

The drift-bias model had the 5 basic parameters plus the drift-bias parameter (v_{bias}). Finally, the hybrid dynamic DDM had 2 biasing parameters (z_{bias} and v_{bias}) and overall 7 free parameters. The diffusion variable was initialized according to equation 8 and evolved in the viewing period according to equation 12 and in the post-stimulus period according to equation 13.

Leaky Accumulator Models – General. The leaky accumulator model was based on models described before [267, 272], constituting an extension of the DDM:

$$dy = (s \cdot v + \lambda \cdot y)dt + cdW \tag{16}$$

where the rate of change of y now also depends on its current value, with a magnitude controlled by the additional parameter λ, the effective leak which reflects the time constant of the accumulation process.

We defined three dynamic variants (c.f. dynamic DDM above) of the leaky accumulator model in order to account for history biases. These different biasing mechanisms were further crossed with two different bound regimes: static or collapsing bounds, as described for the DDM above.

Leaky Accumulator with Starting Point Bias. Here, the diffusion variable was initiated according to equation 8. During the viewing period, it was updated according to:

$$dy(t) = (\lambda \cdot y(t) + g + M_t)dt + cdW \tag{17a}$$

After stimulus offset, accumulation continued according to:

$$dy(t) = \lambda \cdot y(t) + fracy(T)T + cdW \tag{17b}$$

Leaky Accumulator with Input Bias. Here, the diffusion variable was initiated according to equation 6. The evolution of the decision variable during the viewing period was described by:

$$dy(t) = (\lambda \cdot y(t) + g \cdot M_t + v_{bias} \cdot prev)dt + cdW \tag{18}$$

After stimulus offset accumulation continued according to equation 17.2. Responses were determined by a static threshold crossing mechanism, as in the standard DDM models described above.

The third leaky accumulator model we defined, the λ-bias model, accounted for history biases by introducing an asymmetry in the dynamics of evidence accumulation. In this model, we followed a different implementation in order to enable biasing the effective leak (λ) parameter: we reformulated the model to describe two separate accumulators that integrate the sensory evidence. We define the diffusion variable y as $y = y_A y_B$, with y_A and y_B being two independent accumulators coding the right and left choice. The two accumulators were initialized as follows:

$$y_A(0) = U(-sz, sz) \tag{19a}$$
$$y_B(0) = 0 \tag{19b}$$

Starting point variability was thus applied only to one accumulator, which was equivalent to applying this variability on their difference (diffusion variable y).

During the viewing period the two accumulators were updated according to:

$$dy_A(t) = (\lambda_A \cdot y_A(t) + g \cdot f_A(M_t))dt + \frac{cdW}{\sqrt{2}} \tag{20a}$$

$$dy_B(t) = (\lambda_B \cdot y_B(t) + g \cdot f_B(M_t))dt + \frac{cdW}{\sqrt{2}} \tag{20b}$$

The variance of the processing noise applied to each accumulator was divided by 2 such as the processing variance of the accumulators' difference (variable y) is c^2, as in the DDM. The functions f_A and f_B were threshold linear functions with f_A setting negative values to 0 and f_B setting positive values to 0. Specifically:

$$f_A(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{20c}$$

$$f_B(x) = \begin{cases} 0 & \text{if } x > 0 \\ -x & \text{if } x \leq 0 \end{cases} \tag{20d}$$

Thus, the Y_A accumulator "listened" only to the negative values of the input stream while the y_B only to the positive values. The effective leak parameters for each accumulator were defined as follows:

$$\lambda_A = \lambda + f_A(prev) \cdot \lambda_{bias} \tag{20e}$$

$$\lambda_B = \lambda + f_B(prev) \cdot \lambda_{bias} \tag{20f}$$

Leaky Accumulator with Static Bounds. A response was initiated when the difference between the two accumulators (y) exceeded a positive threshold $+a$ (right choice) or surpassed a negative threshold $-a$ (left choice). These leaky accumulator models had 1 biasing parameter each as well as the following 5 basic parameters: threshold value (a), effective leak (λ), noise scaling (), non-decision time (Ter), and starting-point variability (sz).

Leaky Accumulator with Collapsing Bounds. We implemented versions of the leaky accumulator models described above using collapsing bounds. For the input and starting point bias models, the time-varying bounds are described in equations 14.1 and 14.2. For the λ-bias model, collapsing bounds had the same functional form but their asymptote was set to 0 (mirroring the fact that in this model the neutral point of the $y = y_A y_B$ decision variable was at 0, rather than at $a/2$ as in all other models involving a single accumulator):

$$b_{up}(t) = |a - a\frac{t}{t+c}|_0^a \tag{21a}$$

$$b_{down}(t) = |a\frac{t}{t+c} - a|_{-a}^0 \tag{21b}$$

Model fitting procedures We fit the extended models using a Quantile Maximal Likelihood (QMPE) approach. Under this approach, empirical RT values are classified into bins defined by the 0.1, 0.3, 0.5, 0.7 and 0.9 quantiles of the RT distribution (6 bins overall). RT quantiles were derived separately for the various coherence levels. We excluded the 81% coherence trials and pooled together the 0% and 3% coherence trials as RT quantiles in these trials were not distinguishable. This resulted in quantiles for each of 3 difficulty levels (0% and 3%, 9% and 27%), for each of the two responses (correct/ error), and for 2 history conditions (motion direction in current trial consistent or inconsistent with the previous response), leading to 6 bins x 3 coherence x 2 response x 2 history = 72 bins per participant. Denoting the number of empirical observations in a particular bin k by n_k and the probability predicted by the model to derive a response in a particular bin k by P_k, the likelihood L of the data given the model is defined as:

$$L = \prod_k P_k^{n_k} \tag{22}$$

We applied a commonly used multi-stage approach to fit our simulation-based models [299]. First, each fitting session started by generating 20 random parameter sets, drawn from a uniform distribution bounded by the range of each parameter. To improve the precision of likelihood estimates, we generated 10 synthetic trials for each experimental trial, replicating the trials for a given participant. We then computed the likelihood of the model parameters given the data. The parameter set with the best fit out of the initial 20 was used as the starting point for a standard optimization routine ("fminsearchbnd.m" function in Matlab, which implements a constrained version of the Nelder-Mead simplex algorithm). In total, we ran 50 of such fitting sessions, each with a different random seed. Second, we chose the best-fitting parameter set from each of the 50 sessions and recomputed the likelihood while replicating 20 synthetic trials for each experimental trial. Third, the 5 best-fitting of these 50 sets were used as starting points fminsearchbnd, which further refined the local minima of the fit. Fourth, we recalculated the likelihood of the single best parameter set in simulations with 30 synthetic trials for each experimental trial (see eq. 6). For each model, f, AIC values were calculated at the group level:

$$AIC_f = -2\sum_s^n ln(L_s) + 2m_f \tag{23}$$

where N is the total number of participants and s is the participants index. L_s denotes the maximum likelihood estimate for each participant s. Finally, m_f is the number of free parameters for a given model f.

Model simulations We simulated conditional bias functions for various biasing mechanisms within the frameworks of the DDM and the leaky accumulator models. Per biasing mechanism, we simulated 100K traces in time steps of 10 ms using equations 2 (DDM) and 18 (leaky accumulator).

For the DDM simulations (Figure 7 – Figure Supplement 2), the main parameters were: boundary separation = 1; drift rate = 1; non-decision time = 0.1; starting point = 0.5 (expressed as a fraction of the boundary separation); drift bias = 0; drift rate variability = 0.5. We simulated three levels of starting point bias (0.56, 0.62 and 0.68), three levels of constant drift bias (0.2, 0.5 and 0.8), three levels of a time-dependent linear increase in drift bias (1.5/s, 2.5/s and 3.5/s), three levels of constant drift bias (0.2, 0.5 and 0.8) in combination with hyperbolically collapsing bounds (given by equation 16 and using $c = 3$), and three levels of one time-dependent collapsing and one expanding bound: 0.2/s, 0.5/s and 0.8/s.

For the leaky accumulator simulations (Figure 7 – Figure Supplement 3), the main parameters for each accumulator were: input = 1; boundary = 0.42; $\lambda = -2.5$; starting point = 0; input bias = 0. The negative λ's determined that the accumulators were self-excitatory in nature (as opposed to leaky). We choose this to match the primacy effects observed in the data (Figure 7 – Figure Supplement 1d). We simulated three levels of starting point bias (0.05, 0.10 and 0.15), three levels of input bias (0.2, 0.5 and 0.8), and three levels of λ-bias between the two accumulators: (-3 vs -2, -4 vs -1, and -5 vs 0).

We then fit DDM models separately to each of the simulated datasets and fit the parameters boundary separation, drift rate, non-decision time, starting point, drift bias and drift rate variability.

5.4.7 Effective bias signal

We calculated the effective bias signal (as in [273]) for the winning DDM and leaky accumulator models with collapsing bounds (Figure 7d). We focused on those two models because they provided the best fits to the data overall (Figure 7b). We assumed that the current choice is biased in the direction of the previous choice (repetition bias). We arbitrarily set the previous choice to "right" (prev = 1), which means that the biasing mechanisms pushes the decision variable closer to the upper boundary. In both models, the effective bias signal at time t was obtained by dividing the value of the cumulative bias signal by the value of the upper bound on that moment.

For both models, we took the average of the absolute input bias parameter, so as to emulate a repetition bias. For the DDM, we averaged parameters across all participants. For the input bias leaky accumulator model, participants were divided in two groups based on the sign of the fitted parameter λ. Here, we calculated the effective bias signal in two instances: a) by averaging parameters across participants with $\lambda>0$ and b) by averaging parameters across participants with $\lambda<0$. Because the time courses were very similar in these two cases, in Fig. 6d we plotted the average of the two effective bias signals.

For the input bias DDM we computed the cumulative bias signal by taking the difference between the deterministic (i.e., ignoring the noise term and the starting point variability) time course of the decision variable in the input bias model (equation 11) and the deterministic time course of the decision variable in the unbiased model (equation 9). The time course of the upper bound was computed according to

equation 16.1. Correspondingly, for the input bias leaky accumulator model we estimated the cumulative bias signal by subtracting the deterministic decision variable obtained from equation 19.1 from the deterministic decision variable obtained from equation 20. The time course of the upper bound was obtained using equation 22.1.

5.4.8 Statistical tests

We quantified across-subject correlations between P(repeat) and the individual history components in DDM bias parameter estimates using Spearman's rank correlation coefficient. The qualitative pattern of results does not depend on the choice of a specific correlation metric. Even though individual subject parameter estimates are not independent due to the hierarchical nature of the HDDM fit, between-subject variance in parameter point estimates can reliably be correlated to an external variable – in our case, P(repeat) – without inflation of the false positive rate [160]. The difference between two correlation coefficients that shared a common variable, and its associated p-value, was computed using Steiger's test [300].

We used Bayes factors to quantify the strength of evidence across our different datasets. We first computed the Bayes factor for each correlation (between P(repeat) and the history shift in starting point, and between P(repeat) and the history shift in drift bias) [256]. We then multiplied these Bayes factors across datasets to quantify the total evidence in favor or against the null hypothesis of no correlation [257]. BF_{10} quantifies the evidence in favor of the alternative versus the null hypothesis, where $BF_{10} = 1$ indicates inconclusive evidence to draw conclusions from the data. $BF_{10} < 1/10$ or > 10 is taken to indicate substantial evidence for H_0 or H_1 [258].

5.4.9 Data and code availability

All behavioral data, model fits and analysis code are available under a CC-BY 4.0 license at `https://doi.org/10.6084/m9.figshare.7268558`.

5.5 Acknowledgements

We thank Gilles de Hollander and Peter Murphy for discussion. Anke Braun kindly shared behavioral data of the Visual motion 2AFC (FD) study. Christiane Reißmann, Karin Deazle, Samara Green and Lina Zakarauskaite helped with participant recruitment and data acquisition for the Visual motion 2IFC (FD) #2 study.

This research was supported by the German Academic Exchange Service (DAAD, to A.E.U.), the EU's Horizon 2020 research and innovation program (under the Marie Skłodowska-Curie grant agreement No 658581 to K.T.) and the German Research Foundation (DFG) grants DO 1240/2-1, DO 1240/3-1, SFB 936/A7, and SFB 936/Z1 (to T.H.D.). We acknowledge computing resources provided by NWO Physical Sciences.

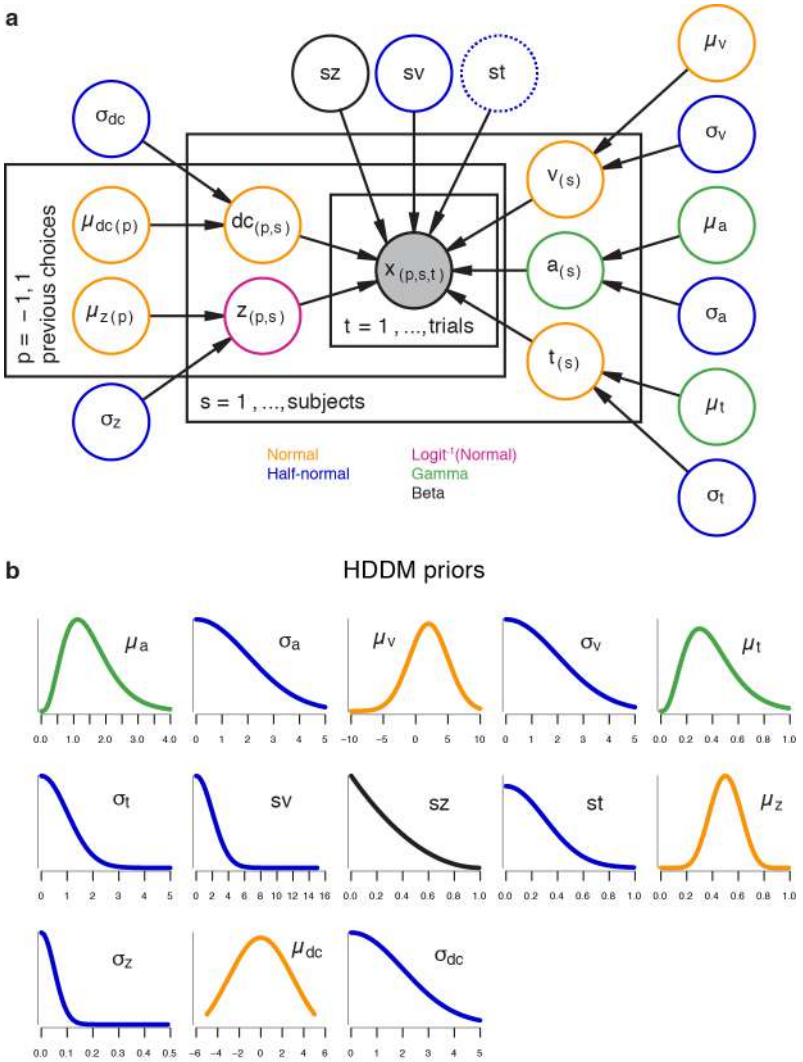


Figure S1: [Figure 3 – figure supplement 1] The hierarchical DDM. (a) Graphical representation of the hierarchical model structure. The full model (with both history-dependent drift bias and starting point) is depicted. Round nodes represent random variables, and the shaded node x represented the observed data (choices and RTs for all observers within each task). Subject-specific parameter estimates were distributed according to the group-level posterior values, thereby 'shrinking' individual values towards the group average. Colors indicate the distributions used for each node. For the datasets with multiple stimulus difficulty levels, we additionally estimated a separate drift rate (v) for each (Figure 3 – Figure Supplement 2a, inset). Between-trial variability in non-decision time was only included in the model shown in Figure 4 – Figure Supplement 1b. **(b)** Prior distributions used for each node, with colors indicating their distribution. See [77] for the full set of prior specifications.

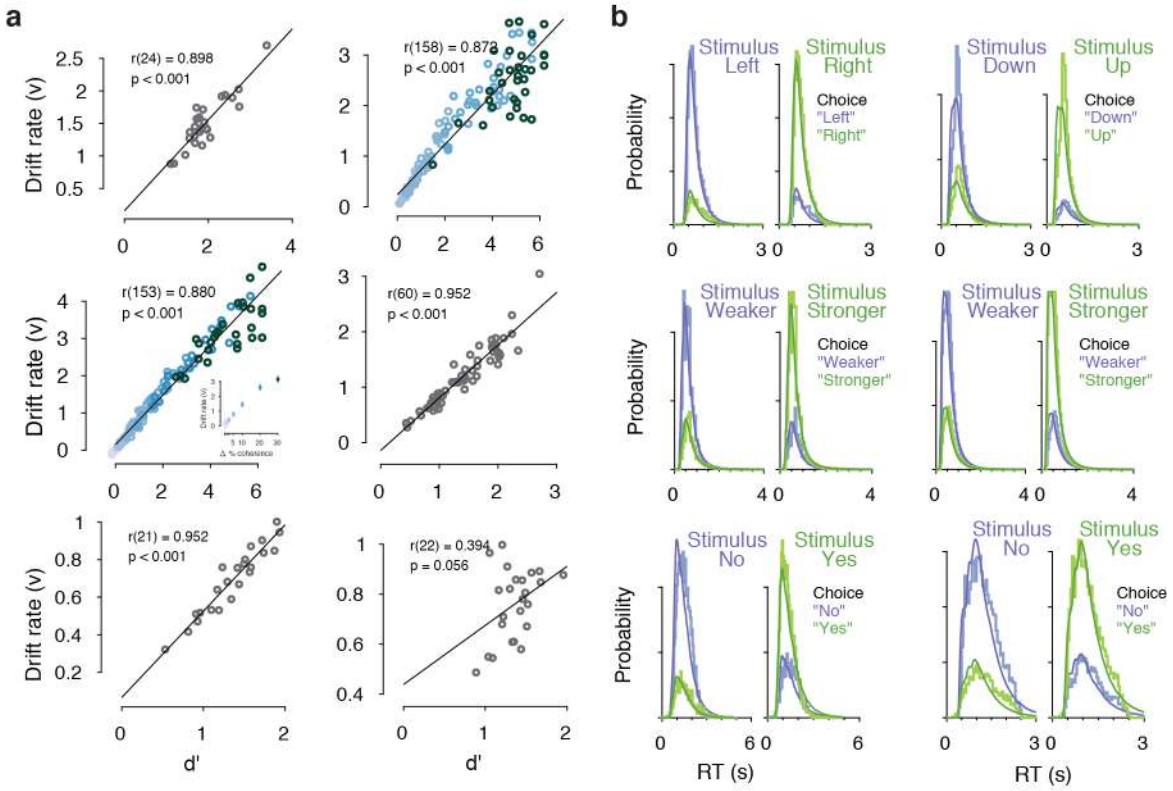


Figure S2: [Figure 3 – figure supplement 2] Drift diffusion model fits. (a) Correlation between drift rate (v) and d'. In the two datasets with multiple stimulus difficulty levels (Visual motion 2AFC (FD) and Visual motion 2IFC (FD) #1), drift rates were estimated separately for each level of stimulus difficulty. In these two datasets, the correlation coefficient displayed is a partial correlation between v and d', while accounting for stimulus difficulty (inset, colors indicate discrete stimulus difficulty levels). As expected, the mean drift rate increased monotonically as a function of evidence strength. **(b)** Measured and predicted RT distributions, across all trials and observers within each dataset. Observed (light) and predicted (dark) RT distributions are separately shown for each combination of stimulus and choice (green/purple), with the low-probability distributions indicating error trials.

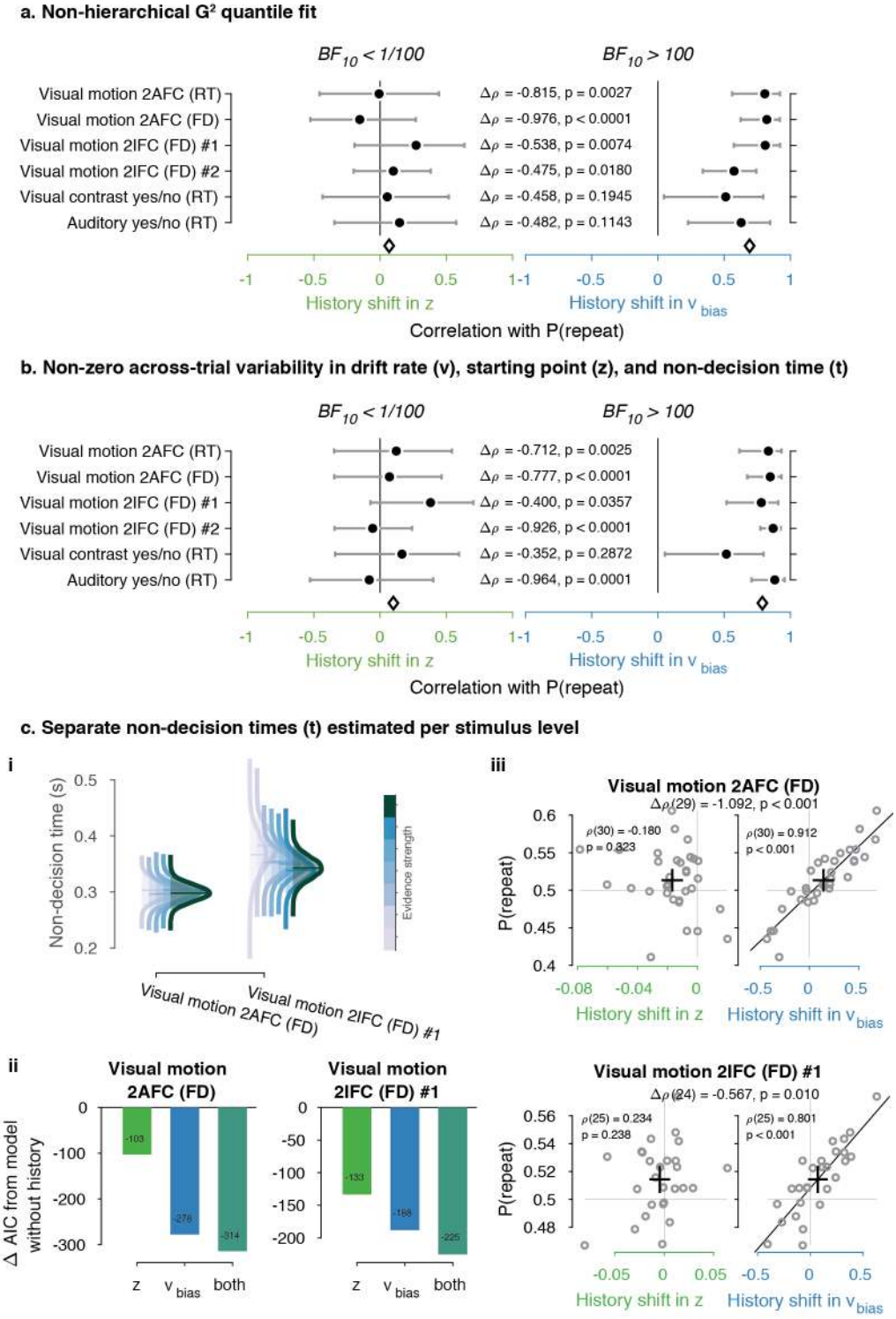

a. Non-hierarchical G² quantile fit
BF10 < 1/100
BF10 > 100
Visual motion 2AFC (RT)
Visual motion 2AFC (FD)
Visual motion 2IFC (FD) #1
Visual motion 2IFC (FD) #2
Visual contrast yes/no (RT)
Auditory yes/no (RT)
Δρ = -0.815, p = 0.0027
Δρ = -0.976, p < 0.0001
Δρ = -0.538, p = 0.0074
Δρ = -0.475, p = 0.0180
Δρ = -0.458, p = 0.1945
Δρ = -0.482, p = 0.1143
History shift in z
History shift in v bias
Correlation with P(repeat)
b. Non-zero across-trial variability in drift rate (v), starting point (z), and non-decision time (t)
Δρ = -0.712, p = 0.0025
Δρ = -0.777, p < 0.0001
Δρ = -0.400, p = 0.0357
Δρ = -0.926, p < 0.0001
Δρ = -0.352, p = 0.2872
Δρ = -0.964, p = 0.0001
c. Separate non-decision times (t) estimated per stimulus level
i
Non-decision time (s)
Evidence strength
Visual motion 2AFC (FD)
Visual motion 2IFC (FD) #1
ii
Visual motion 2AFC (FD)
Visual motion 2IFC (FD) #1
Δ AIC from model without history
z
v bias
both
iii
Visual motion 2AFC (FD)
Δρ(29) = -1.092, p < 0.001
ρ(30) = -0.180
p = 0.323
ρ(30) = 0.912
p < 0.001
P(repeat)
Visual motion 2IFC (FD) #1
Δρ(24) = -0.567, p = 0.010
ρ(25) = 0.234
p = 0.238
ρ(25) = 0.801
p < 0.001

Figure S3: [Figure 4 – figure supplement 1] Control model fits. (a) Summary figure based on non-hierarchical G2 fits [298]. Rather than the full RT distribution, we fit each observers' RT quantiles (0.1, 0.3, 0.5, 0.7, 0.9) and correlated history shifts in the DDM bias parameters to individual P(repeat), as in Figure 4. **(b)** Summary figure based on the full hierarchical model, where across-trial variability in non-decision time (st) was added as a free parameter. Like the across-trial variability in drift rate (sv) and starting point (sz), the st parameter was only estimated at the group level [112]. **(c)** Fits of the two datasets with multiple evidence strengths, allowing non-decision time to vary with these discrete levels (see Methods). (i) Posterior distributions of group-level non-decision time for each level of evidence strength. (ii) Comparison between models accounting for choice history bias through a starting point, drift bias or both, while allowing for non-decision time to vary with levels of stimulus strength. (iii) Correlations between individual history shifts in starting point or drift bias and repetition probability, from a model where non-decision time was separately estimated per level of stimulus strength.

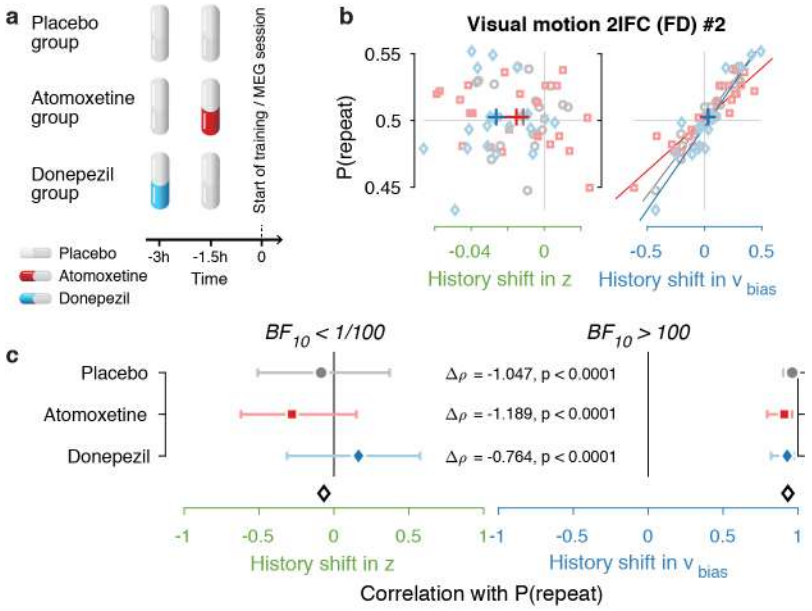


Figure S4: [Figure 4 – figure supplement 2] Same biasing mechanism under two pharmacological interventions. (a) Participants in the MEG study were assigned to one of three pharmacological groups. At the start of each experimental session, they orally took 40 mg atomoxetine (Strattera®), 5 mg donepezil (Aricept®), or placebo. Since the time of peak plasma concentration is 3 hours for donepezil [301] and 1-2 hours for atomoxetine [302], we used a placebo-controlled, double-blind, double-dummy design, entailing an identical number of pills at the same times before every session for all participants. Participants in the donepezil group took 5 mg of donepezil 3 hours, and placebo 1.5 hours before starting the experimental session. Participants in the atomoxetine group took placebo 3 hours, and 40 mg of atomoxetine 1.5 hours before the experimental session. Those in the placebo group took identical-looking sugar capsules both 3 and 1.5 hours before starting the session. This ensured that either drug reached its peak plasma concentration at the start of the experimental training. The drug doses were based previous studies with healthy participants [303, 304]. Blood pressure and heart rate were measured and registered before subjects took their first and second pill. In the three hours before any MEG or training session, participants waited in a quiet room. In total, 19 people in the placebo, 22 in the atomoxetine, and 20 in the donepezil group completed the full study. **(b, c)** Choice history biases separately for each pharmacological group. Since we did not observe differences in choice history bias between these groups, we pooled all observers for the main analyses.

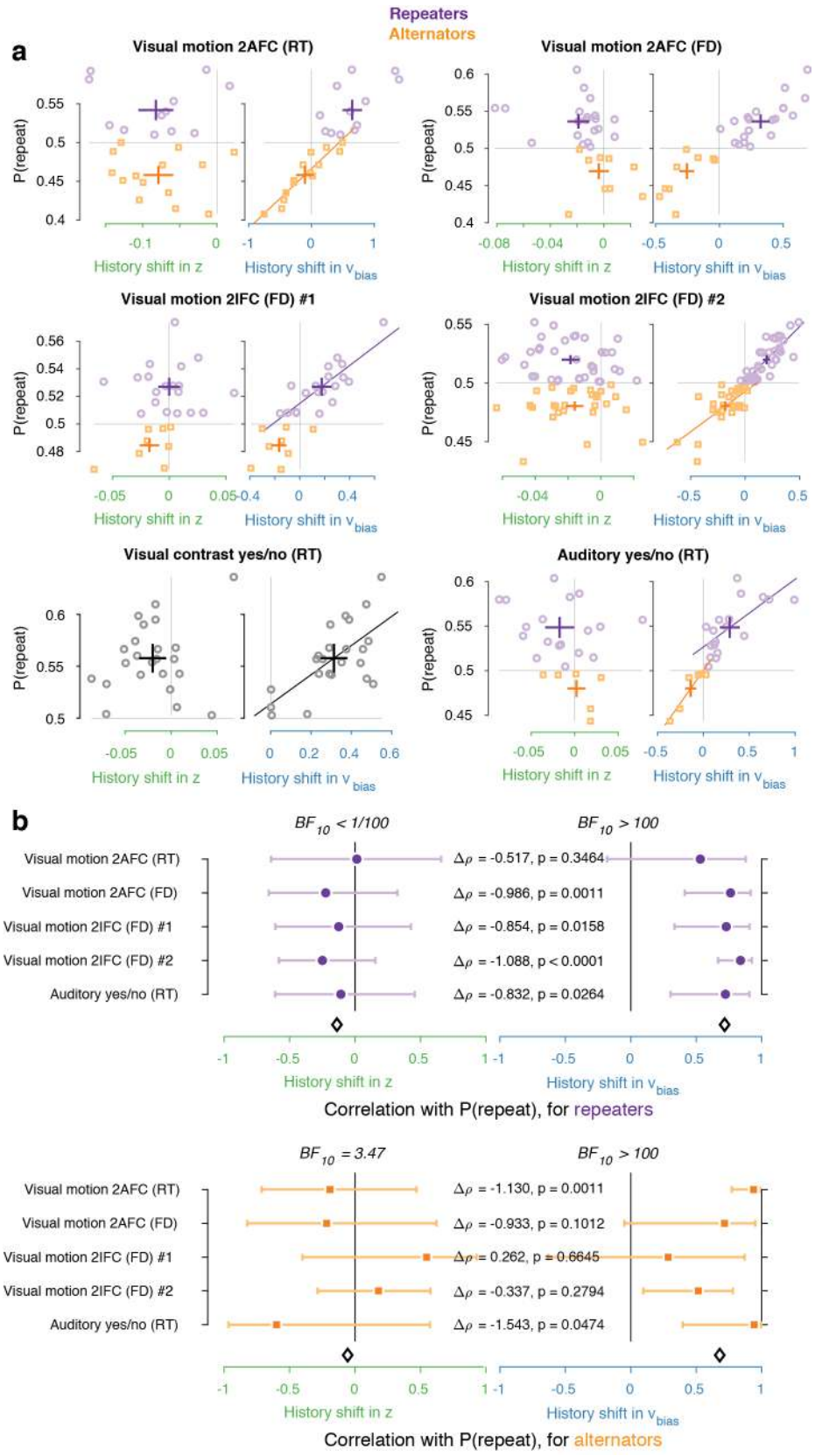

a
Repeaters
Alternators
Visual motion 2AFC (RT)
Visual motion 2AFC (FD)
Visual motion 2IFC (FD) #1
Visual motion 2IFC (FD) #2
Visual contrast yes/no (RT)
Auditory yes/no (RT)
P(repeat)
History shift in z
History shift in v bias
b
BF10 < 1/100
BF10 > 100
Visual motion 2AFC (RT)
Visual motion 2AFC (FD)
Visual motion 2IFC (FD) #1
Visual motion 2IFC (FD) #2
Auditory yes/no (RT)
Δρ = -0.517, p = 0.3464
Δρ = -0.986, p = 0.0011
Δρ = -0.854, p = 0.0158
Δρ = -1.088, p < 0.0001
Δρ = -0.832, p = 0.0264
Correlation with P(repeat), for repeaters
BF10 = 3.47
BF10 > 100
Δρ = -1.130, p = 0.0011
Δρ = -0.933, p = 0.1012
Δρ = 0.262, p = 0.6645
Δρ = -0.337, p = 0.2794
Δρ = -1.543, p = 0.0474
Correlation with P(repeat), for alternators

Figure S5: [Figure 4 – figure supplement 3] Repeaters vs. alternators. (a) Correlations between individual history shifts in starting point or drift bias and repetition probability, fit separately for groups of 'repeaters' (P(repeat) > 0.5) and 'alternators' (P(repeat) < 0.5). **(b)** Summary of correlations (as in Figure 4c) separately for repeaters and alternators. Error bars indicate the 95% confidence interval of the correlation coefficient. $\Delta\rho$ quantifies the degree to which the two DDM parameters are differentially able to predict individual choice repetition, p-values from Steiger's test. The black diamond indicates the mean correlation coefficient across datasets. In the Visual contrast yes/no RT dataset there were no alternators, so this dataset was excluded from the overview.

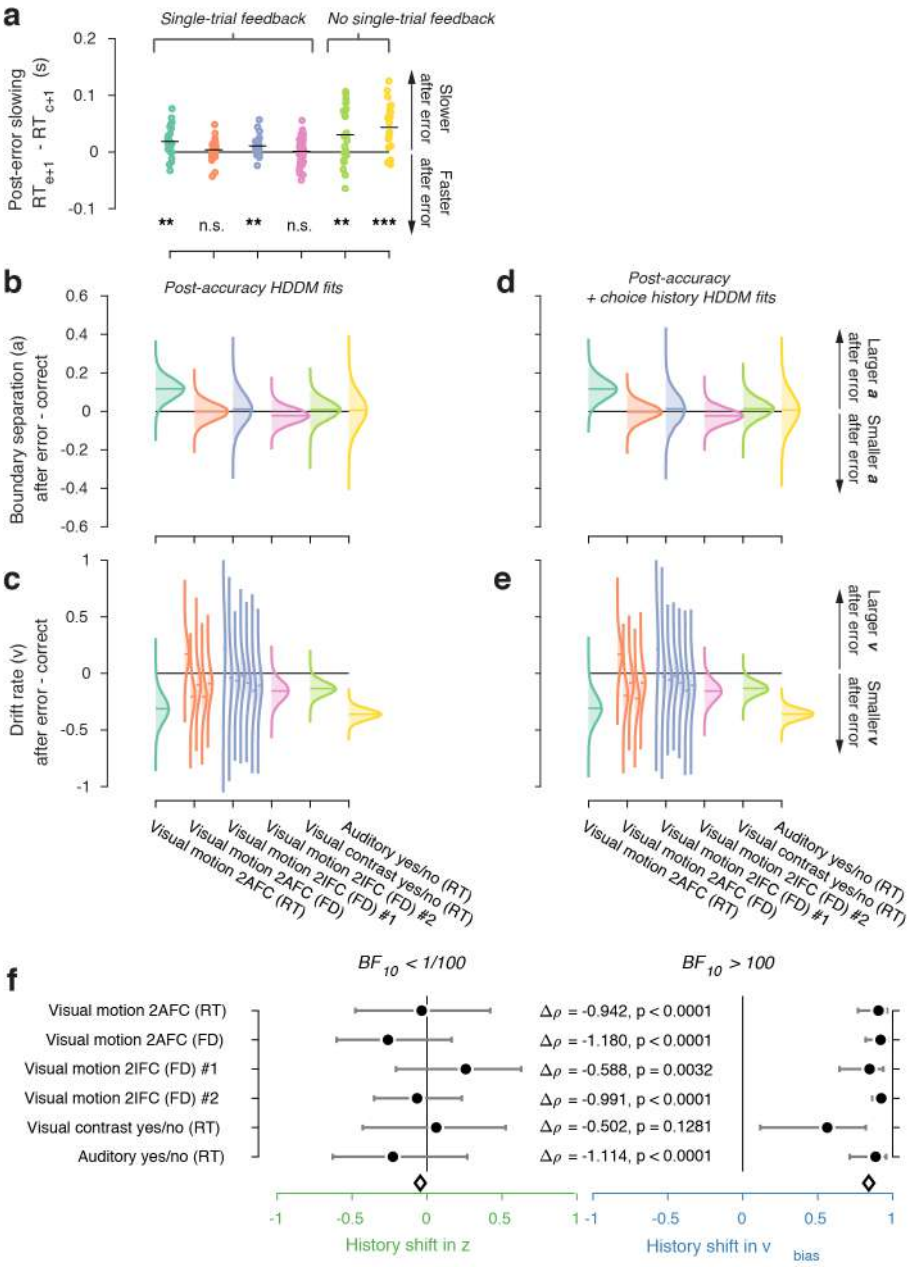


Figure S6: [Figure 4 – figure supplement 4] Post-error slowing. (a) We computed post-error slowing as the difference in average RT after error and correct trials. In four of the six datasets, observers showed significant post-error slowing (permutation test against zero; *** $p < 0.001$, ** $p < 0.01$, n.s. $p > 0.05$). We then fit an HDDM model where the overall drift rate, as well as the boundary separation, were allowed to vary depending on the outcome (correct vs. error) of the previous trial. **(b)** Changes in boundary separation after error vs. correct trials. **(c)** Changes in drift rate after error vs. correct trials. For the two datasets with multiple levels of stimulus strength, the effect of previous error vs. correct on drift rate is shown separately for each discrete level of stimulus strength (weak to strong evidence from left to right; see also Methods and Figure 3 – Figure Supplement 2a). **(d-e)** as (b-c), but from a HDDM model where we simultaneously allowed choice history to affect starting point and drift bias, as well as previous accuracy to affect boundary separation and drift rate. Distributions were smoothed using kernel density fits. Shaded regions represent the 95% BCI, and white lines indicate the posterior mean. Errors were succeeded by a decrease in mean drift rate in most datasets and by an increased boundary separation in some datasets: both effects conspired to slow down decisions after an error. **(f)** Correlations between P(repeat) and history shifts in starting point and drift bias, from the model where previous accuracy was also allowed to affect boundary separation and drift rate.

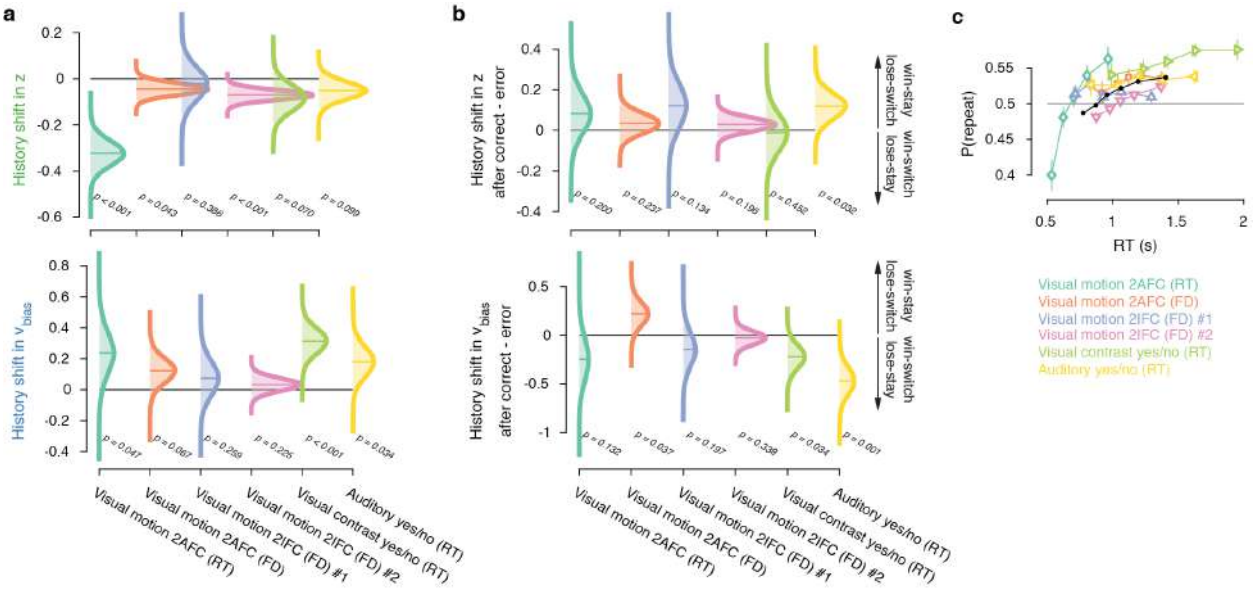


Figure S7: [Figure 4 – figure supplement 5] Group-level posterior distributions of history bias parameters. History shift in **(a)** starting point and drift bias, as well as the effect of previous choice outcome on **(b)** starting point and drift bias, separately for each dataset. The posteriors were taken from the model where history-dependent z and vbias were fit simultaneously. Distributions were smoothed using kernel density fits. Shaded regions represent the 95% BCI, and white lines indicate the posterior mean. P-values indicate the overlap with zero (a), or between the two conditions (b). **(c)** Repetition probability for five quantiles of the RT distribution, separately for each dataset (colors) and averaged across datasets (black). For those RTs < 600ms, group-level behavior tended towards choice alternation.

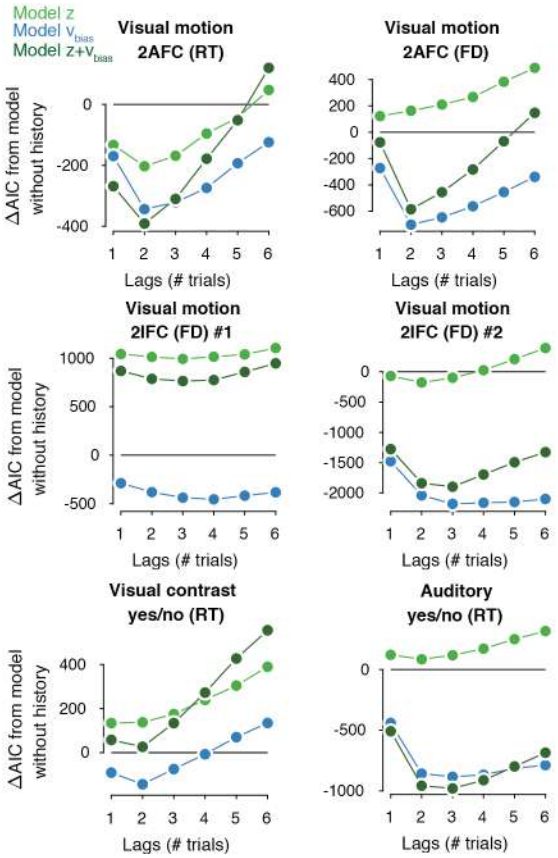


Figure S8: [Figure 6 – figure supplement 1] Contribution of previous choices to current drift and starting point bias as function of lag. Model comparison (Δ AIC from a baseline model without history) between models where previous correct and incorrect choices affected only starting point (light green squares), only drift bias (blue diamonds), or both (blue-green circles), up to 6 past trials. The best-fitting model is indicated by a black outline.

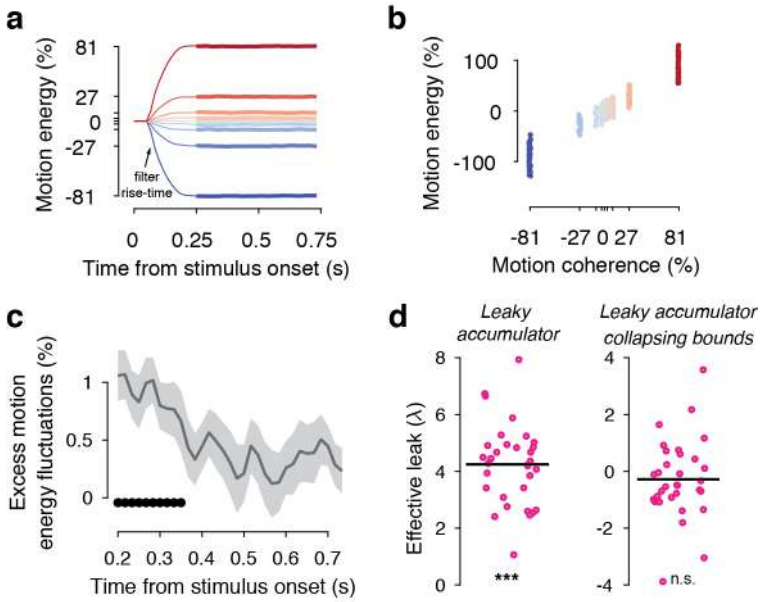


Figure S9: [Figure 7 – figure supplement 1] Motion energy filtering, psychophysical kernels and the effective time-constant of evidence integration. **(a)** Within each generative coherence level, the average motion energy traces were rescaled to express motion energy in % coherence. The initial 200ms of the trial fall within the rise-time of the spatiotemporal motion energy filter. **(b)** The average motion energy is a linear function of coherence, with substantial trial-by-trial fluctuations that arise from the stochastically generated noise dots. Note that while we previously used motion energy filtering on the dot coordinates in the Visual motion 2IFC (FD) #1 dataset (Urai et al., 2017), the large stimulus display in that study resulted in most noise fluctuations being averaged out over space. This resulted in only small trial-by-trial differences in the effective decision-relevant input). **(c)** Psychophysical kernels, indicating the effect of fluctuations in motion energy (using the three weakest coherence levels) on observers' choice over time. Shaded errorbars indicate group s.e.m., black dots show significant ($p < 0.05$, FDR-corrected) group-level deviations from zero. **(d)** Individual effective leak parameters, estimated from an leaky accumulator model either without (left) or with (right) collapsing. A positive λ indicates that the accumulators accelerate towards the decision bound, depending on the value of the decision variable. When including a collapsing bound in the model, which captures an overall urgency in the decision process, overall effective leak biases are no longer significantly different from zero. *** $p < 0.001$, ** $p < 0.01$, n.s. $p > 0.05$, permutation test against zero.

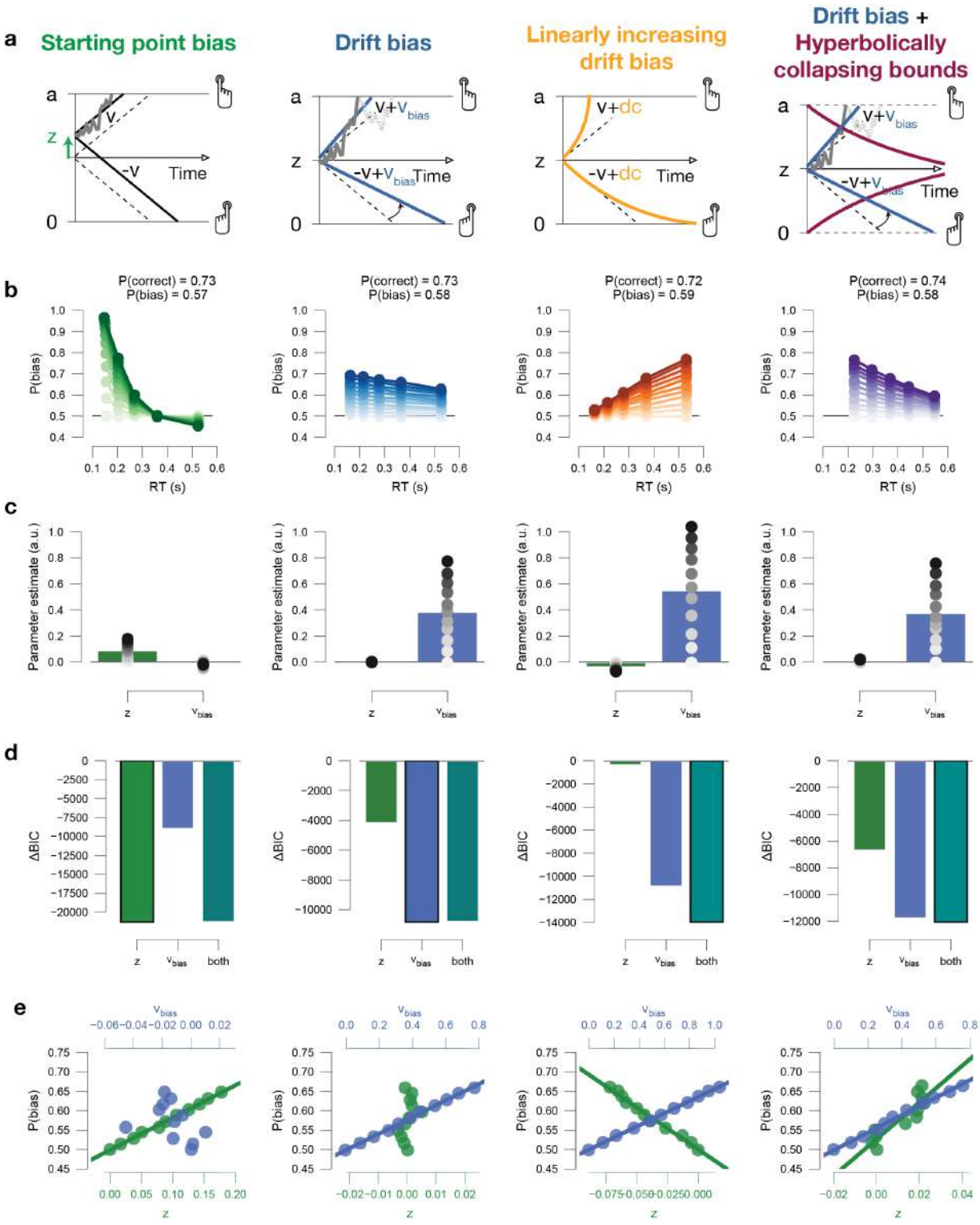

a
Starting point bias
Drift bias
Linearly increasing drift bias
Drift bias + Hyperbolically collapsing bounds
a
z
0
v
-v
Time
v+V bias
-v+V bias
v+dc
-v+dc
b
P(correct) = 0.73
P(bias) = 0.57
P(correct) = 0.73
P(bias) = 0.58
P(correct) = 0.72
P(bias) = 0.59
P(correct) = 0.74
P(bias) = 0.58
P(bias)
RT (s)
c
Parameter estimate (a.u.)
z
V bias
d
ΔBIC
z
V bias
both
e
V bias
P(bias)
z

Figure S10: [Figure 7 – figure supplement 2] Drift diffusion model simulations. (a) Model schematics. **(b)** Conditional bias functions, for synthetic individuals with increasing bias magnitude (indicated by color intensity). **(c)** Parameter estimates of a hybrid HDDM (with both starting point and drift bias) fit on simulated data. **(d)** Model comparison from a baseline model without history. The best-fitting model (lowest BIC) is indicated by a black outline. **(e)** Correlations between P(bias) of each synthetic individual, and their respective parameter estimates for both starting point and drift bias.

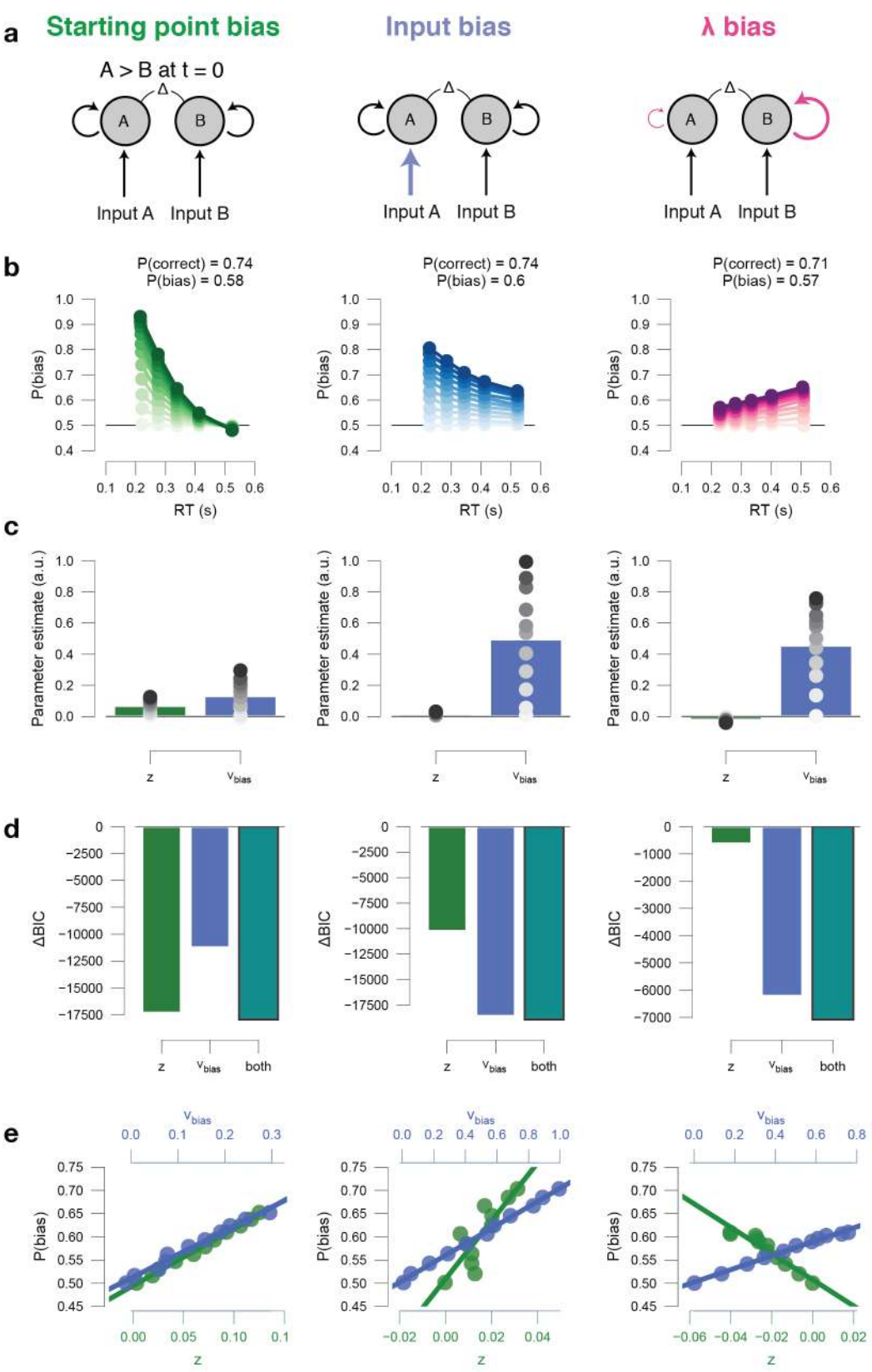

a
Starting point bias
Input bias
λ bias
A > B at t = 0
Δ
A
B
Input A
Input B
b
P(correct) = 0.74
P(bias) = 0.58
P(correct) = 0.74
P(bias) = 0.6
P(correct) = 0.71
P(bias) = 0.57
P(bias)
RT (s)
c
Parameter estimate (a.u.)
z
v_bias
d
ΔBIC
z
v_bias
both
e
v_bias
P(bias)
z

Figure S11: [Figure 7 – figure supplement 3] Leaky accumulator model simulations. To match the empirical data (Figure 7 – Figure Supplement 1d), we choose the overall leak parameter λ to be > 0, producing a primacy effect through self-excitatory accumulators, whereby evidence early on in the trials has a stronger leverage on choice. We verified that the same conclusions hold for $\lambda \leq 0$ (data not shown). All simulations included collapsing bounds. **(a)** Model schematics. Left to right: (i) The effect of a biased starting point on choice declines rapidly with elapsed time, as the decision variable is increasingly governed by the input. (ii) The effect of biased input to the accumulation stage within the leaky accumulator model is uniquely captured by drift bias within the DDM. Also, the effect input bias on choice increases with elapsed time, as the decision variable is increasingly governed by the λ_{bias}. (iii) The effect of biased leak of the accumulators (henceforth termed 'λ_{bias}') on choice increases with elapsed time, as the decision variable is increasingly governed by the λ_{bias}. **(b)** Conditional bias functions, for synthetic individuals with increasing bias magnitude (indicated by color intensity). **(c)** Parameter estimates of a hybrid HDDM (with both starting point and drift bias) fit on simulated data. **(d)** Model comparison from a baseline model without history. The best-fitting model is indicated with a black outline. **(e)** Correlations between P(bias) of each synthetic individual, and their respective parameter estimates for both starting point and drift bias.
It noteworthy that λ_{bias} in the leaky accumulator model predicts the strongest choice bias for long RTs, while both DDM starting point and drift bias predict the strongest choice bias for short RTs. This implies that if a major source of choice bias in any dataset is due to a leak bias, the DDM is not going to be able to easily account for this. Our simulations show that the best-fitting DDM such data shows: (i) a drift bias, in order to explain the choice bias for relatively long RTs, and (ii) a starting point of opposite sign, in order to push down the expected choice bias for relatively short RTs. Indeed, these opposing effects of starting point and drift bias can be observed in the bars in (h), and are present in some of the datasets used here (Figure 4 – Figure Supplement 5a). The present simulation results suggest that even stronger choice repetition (or alternation) effects as measured here would give rise to opposite effects on starting point and drift bias in DDM fits.

A

Summary

A.1 English

Many decisions are made in the face of uncertainty and involve accumulating ambiguous evidence over time. This holds for elementary decisions based on perceptual evidence (finding your way home through a rainy windshield) ranging to more complex decisions based on abstract evidence (deciding which stock to buy, or which career to pursue). Protracted decision-making is implemented by a distributed brain network: in the context of a perceptual decision, sensory cortex encodes input signals ("evidence samples"), association cortex accumulates these over time into a decision variable, and motor regions translate the decision into a behavioral act. Strikingly, when asked to make repeated decisions, we often choose differently each time, even when we are given the same information to inform our choice. This implies that the path a decision-making process takes in our brain is not fixed, but, to an extend, variable. Behavioral variability is commonly attributed to intrinsic noise in the nervous system.

However, behavioral variability is perhaps not completely unpredictable: a substantial part of the variability in behavior might in fact be explained and predicted by tracking the momentary "arousal state" of the brain, and by tracking the "history of choices" a particular decision is embedded in. Indeed, the brain's arousal state is briefly elevated during decision-making. The arousal state is regulated by brainstem nulei, which have widespread projections that release modulatory neurotransmittors (neuromodulators). Neuromodulators alter the operating mode of their target circuits, for example, by regulating their overall excitability. As such, they determine the arousal state of the brain and tune information processing. Also, choices are rarely made in isolation, but often depend on a person's previous experiences (when finding your way home through that rainy windshield, you probably don't want to make two right-turn decisions in a row; unless you're sure you went the wrong way). This implies that previous choices can bias the brain's decision-making machinery. Taken together, the brain's arousal state and the history of

choices should, *somehow*, shape cortical decision-computations and overt choices, and thereby explain a significant fraction of the observed behavioral variability. The goal of this thesis is to determine *how*.

Here, this question is addressed with a series of experiments combining psychophysics, computational modeling, pupillometry, functional magnetic resonance imaging (fMRI) of the brainstem, and magnetoencephalography (MEG) of cortical decision computations. The first chapter establishes that pupil size at constant luminance is a reliable measure of the activity of a number of neuromodulatory brainstem nuclei known to regulate cortical arousal. Further, the first chapter demonstrates that the functional impact of pupil-linked phasic arousal is to suppress suboptimal biases in evidence accumulation and subsequent choice behavior. The second chapter establishes that the arousal-linked suppression of suboptimal biases in evidence accumulation is a general principle: it occurs in a range of decision-making tasks (when decisions are based on perceptual evidence, are made from memory, or based on numerical values) and across species (humans and mice). The third chapter demonstrates that the pupil-linked bias suppression (in a perceptual decision-making context) is mediated by the arousal-related shaping of choice-predictive alpha-band activity (a signature of "excitability") in regions along the dorsal stream of the visual cortical hierarchy, including primary visual cortex. The fourth chapter establishes that pupil-linked phasic arousal is recruited in relation to key computational variables such as uncertainty and surprise. The fifth and last chapter highlights that the evidence accumulation process is not only driven by the momentary evidence, bias, and arousal, but indeed also by previous choices.

Thus, phasic arousal and previous choices calibrate a key computation during decision-making: evidence accumulation. When decisions are made in the face of uncertainty tracking phasic arousal signals and the history of choices may be just as important for predicting choice behavior as tracking the objective evidence gathered from the outside world. This thesis therefore establishes that phasic arousal signals and the history of choices account for a substantial fraction of the apparent variability in overt behavior.

A.2 German

Wenn wir Entscheidungen treffen, so tun wir dies oftmals auf basierend auf mehrdeutiger oder unzuverlässiger Information, welche über die Zeit akkumuliert wird. Dieser Prozess der Informationsakkumulation lässt sich sowohl bei sehr einfachen perzeptuellen Entscheidungen, das heißt Entscheidungen, die auf rein sensorischer Information basieren (z.B. die Detektion von Hindernissen auf der Fahrbahn durch eine verregnete Frontscheibe), jedoch auch bei komplexeren Entscheidungen (z.B. ob eine Aktie gekauft, oder ein bestimmter Karriereweg einschlagen werden soll) beobachten. Im Gehirn sind solche zeitlich ausgedehnten Entscheidungsprozesse in verteilten neuronalen Netzen implementiert: bei perzeptuellen Entscheidungen enkodieren sensorische kortikale Areale die eingehende sensorische Information, wohingegen höhere Areale des Assoziationskortex diese Information über die Zeit, in Form einer Entscheidungsvariable, akkumulieren. Regionen des motorischen Kortex hingegen transformieren die akkumulierte Information in einen behavioralen Output. Interessanterweise zeigt sich in Situationen, in welchen Menschen wiederholt Entscheidungen basierend auf sehr ähnlicher oder gar identischer Information treffen

müssen, dass die getroffenen Entscheidungen variabel sind – selbst dann, wenn die zugrundeliegende Information bei jeder dieser Entscheidungen identisch ist. Dies suggeriert, dass der Prozess der Entscheidungsfindung nicht vollends deterministisch verläuft, sondern vielmehr einer gewissen Variabilität unterliegt. Diese Variabilität im Entscheidungsprozess wird häufig intrinsischem Rauschen in neuronalen Netzen zugeschrieben. Intrinsisches Rauschen erklärt jedoch nur einen Teil der resultierenden behavioralen Variabilität. Ein weiterer Teil kann einerseits durch Schwankungen im allgemeinen Erregtheitszustand des Gehirns erklärt werden, andererseits durch die Berücksichtigung vergangener Entscheidungen auf welche eine gegenwärtige Entscheidung folgt.

Im Verlauf des Entscheidungsprozesses werden neuromodulatorische Systeme des Hirnstamms aktiviert. Diese Systeme verfügen über weitreichende anatomische Verbindungen in den Neokortex, welche Neuromodulatoren in die Zielregion freigeben. Neuromodulatoren verändern den Operationsmodus neuronaler Schaltkreise, zum Beispiel durch die Regulation der neuronalen Erregbarkeit. Durch ihre weitreichenden Verbindungen beeinflussen sie den globalen Erregtheitszustand des Gehirns und folglich die neuronale Informationsverarbeitung. Entscheidungen werden zudem selten in vollständiger Isolation getroffen, sondern sind vielmehr eingebettet in eine Abfolge vieler Entscheidungsprozesse, welche ebenfalls Einfluss auf den neuronalen Entscheidungsprozess nehmen. Zusammengenommen prägen der momentane Erregtheitszustand des Gehirns sowie die genaue Abfolge an vorhergehenden Entscheidungen die Informationsverarbeitungsprozesse welche der Entscheidungsfindung zugrunde liegen. Variabilität im Erregtheitszustand oder in der Entscheidungsgeschichte erklären damit einen bedeutenden Teil beobachteter Verhaltensvariabilität - selbst angesichts identischer Information. Das Ziel der vorliegenden Dissertation ist es darzulegen, über welche neuronalen Prozesse diese Prozesse implementiert sind.

Diese Frage wird hier durch eine Kombination aus Psychophysik, mathematischer Modellierung, sowie funktioneller Magnetresonanztomographie und Magnetoenzephalographie adressiert. Im ersten Kapitel wird der Pupillendurchmesser (bei konstanter Helligkeit) als ein Marker für die Aktivität neuromodulatorischer Nuklei des Hirnstamms, und damit für globale Erregung, etabliert. Weiterhin wird im ersten Kapitel gezeigt, dass in Phasen erhöhter Erregung suboptimale Biases in der perzeptuellen Entscheidungsfindung reduziert bzw. unterdrückt sind. Im zweiten Kapitel wird die Frage nach der Allgemeingültigkeit der genannten Reduktion suboptimaler Biases behandelt. Es wird dargelegt, dass dieses Prinzip einerseits bei einer Vielzahl unterschiedlicher Entscheidungsaufgaben variierender Komplexität, sowie in unterschiedlichen Spezies (Menschen und Mäusen) auftritt. Im dritten Kapitel wird dargelegt, dass die genannte Unterdrückung suboptimaler Biases durch erregungsbedingte Einflüsse auf Alpha-Band Aktivität, entlang des dorsalen Stroms der visuellen Hierarchie, mediiert wird. Im vierten Kapitel wird gezeigt, dass phasische Erregung in Relation zu Schlüsselkomponenten der Informationsverarbeitung, wie z.B. Sicherheit/Unsicherheit und Überraschung, rekrutiert wird. Im fünften und letzten Kapitel wird aufgezeigt, dass Informationsakkumulation nicht nur von gegenwärtiger Information, etwaigen Biases sowie allgemeiner Erregung abhängt, sondern zudem von der Reihenfolge an vorhergehenden Entscheidungen beeinflusst wird.

Zusammenfassend wird in dieser Dissertation aufgezeigt, wie neuronale Erregung sowie vorhergehende Entscheidungensequenzen Informationsverarbeitung während unterschiedlicher Entscheidungsaufgaben beeinflussen. In Situationen, in welchen Entscheidungen basierend auf unzuverlässiger oder ambivalenter externer Information getroffen werden, haben Faktoren wie neuronale/psychische Erregung, sowie die Sequenz vorhergehender Entscheidungen, einen ähnlich großen Einfluss auf künftige Entscheidungen wie die tatsächliche und objektive Information, welche vom Entscheider aus der Umwelt extrahiert wird. Dabei zeigt sich, dass nur ein Teil der beobachteten Variabilität im Entscheidungsverhalten intrinsischem neuronalem Rauschen zugeschrieben werden kann.

B

Acknowledgements

I have so many amazing people around me, and I want to thank everyone for who they are and for making my life meaningful. In the context of this thesis, there are some that I would like to mention in more detail, and I will attempt to do so in chronological order.

I want to thank my mom and dad for providing a secure base from which I could explore the world and where I continue to come back to, for not pressuring me to grow up too fast, for showing me the highest mountains and deepest oceans, and for always believing that I knew best what was right for me (but also for asking the occasional critical question). I want to thank my big sister for paving the road in so many ways and for being my best friend ever since I opened my eyes – I will always look up to you. I want to thank the rest of my amazing family, including the Mahamuni's, for their support, interest, and very valid questions regarding the practical implications of what I'm doing.

I want to thank Tobias for being my mentor for almost a decade. I'm proud to say that my scientific thinking is largely shaped by you: after all, you did not only supervise this PhD thesis, but also guided me through both my BSc and MSc graduations. I still often think "what would Tobi do", and then the solution is never far away. The combination of your breadth of knowledge, clear thinking and genuine excitement about unexpected outcomes was a great resource to me, and above all incredibly inspiring. I fondly remember the bike rides through Brooklyn, early mornings at the Golden Pudel in Hamburg, and the walks along the Amstel river in Amsterdam.

I want to thank Tomas for passing on so much of his skill- and tool-sets. Learning eye-tracking, fMRI, and programming (in Python!) from you greatly enabled me as a young researcher. Above all, your endless enthusiasm about new experiments was very contagious: science is just so much fun! I won't forget us working together from your house in de Nieuwe Oostenburgerstraat, trying to squeeze pupil data through your deconvolution scripts, while eating home-cooked Thai soup or the best fried egg I have ever had.

I want to thank my high-school buddies and roommates Daan and Remy for being the very first participants in my first experiment. With your help, and we're talking about +40 hours of staring at screen that was mostly just grey, I was able to fine-tune the experimental design that I have used throughout my PhD track. Here's to many more coffees and beers in life! Here I also want to take the opportunity to thank all the subjects that have participated in my experiments over the years.

I want to thank everyone in the programme group Brain and Cognition at the University of Amsterdam for being such a warm and fun crowd. I vividly remember the "lunch, lunch, lunch" cries, and everyone cramming around the small table. Specifically, I want to thank Niels and Simon for accepting me in their cool-kids' room when we were still in de Diamantbeurs. The loud (and pretty appalling) music at the end of the afternoon was always a welcome reminder to not get too worked up. Daniel should also be mentioned here – those ping-pong sessions in the summer of 2014 were quite something.

I want to thank each and every member of the Donnerlab for all the good times in and outside Hamburg: Anne, for being both the moral and practical compass of the lab – without you, I would never have heard of open science, bioRxiv, Twitter, Slack, Zotero, I would have missed many deadlines, and definitely would not have had such a massive last chapter of my thesis; Thomas, for the epic office boys picture, for refreshing conversations about life after science and computers, and for equally refreshing drinks on your bankje in de voortuin; Olympia, for making the long days of MRI scanning very bearable, for being the social glue in the lab, and for renting Srushti and me your beautiful house; Niklas, for being so well-considered and always supportive – I'm looking forward to the next occasion talking about the ups and down of life and drinking beer on a jetty; Tom and Rudy, for displaying such a charming combination of rigorous science and childish behavior; Tom once more for translating the summary of this thesis into German, and also for obtaining a proof of residence that I needed last minute to get married – in short, without you I would be single and without a degree; Peter, for being Dr. Murphy – my biggest regret in life so far is not having published with you, but no one will ever take away the fact that my desk was directly facing and touching yours; Konstantinos, for your good sense of humor and rad modeling skills; Kobe, for being such a competitive and elegant ping-pong player; Bharath, for your Indian hospitality and impressive childhood stories that often involved extreme sports (kabaddi) (PS: I'm so happy for you and Priya!); Anke, for introducing us to Hafenklang and headCRASH; Keno, for always being happy despite the various hangovers and rugby injuries, and for making the ice-skating trip to Appelbeke happen; Christoffer, for giving me a glimpse into the world of dungeons and dragons (once was enough), and for knowing all the Hamburg ins and outs, Alena, for the amazing cocktails, although it's going to be a while before I try a Mai Tai again; Carolina for brining that Southern vibe to rainy Hamburg!

I want to thank Andreas Engel and the rest of the department of Neurophysiology at the University Medical Center Hamburg-Eppendorf for all the support, and Philipp Sterzer and Christian Büchel for evaluating my work.

I want to thank Matthew McGinley for the cool cross-species transatlantic collaboration that started years ago. This culminated in one of the chapters presented here, and in a postdoc position in his lab

at Baylor College of Medicine in Houston. I'm sure we're going to do many more great things together, because we already are.

Finally, I want to thank Srushti for being the love of my life, for moving continents with me, for being so involved and interested, for keeping me sane, for sharing me with my "2nd girlfriend" called Tobias, and for so many more things. Above all, I learned from you to fully commit. I gained so much more from this than I thought possible, this thesis being only a modest fraction of that.

C

Curriculum Vitae

Lebenslauf entfällt aus datenschutzrechtlichen Gründen.

D

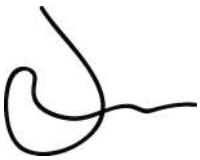
Eidesstattliche Erklärung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift:



References

[1] Paul W Glimcher. Indeterminacy in brain and behavior. *Annual review of psychology*, 56:25–56, 2005.

[2] Joshua I Gold and Michael N Shadlen. The neural basis of decision making. *Annual review of neuroscience*, 30:535–574, 2007.

[3] M. N. Shadlen, K. H. Britten, W. T. Newsome, and J. A. Movshon. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 16(4):1486–1510, February 1996. ISSN 0270-6474.

[4] Leo P Sugrue, Greg S Corrado, and William T Newsome. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, 6(5):363–375, May 2005.

[5] V Wyart and E Koechlin. Choice variability and suboptimality in uncertain environments. *Current Opinion in Behavioral Sciences*, 2016.

[6] G Aston-Jones and J D Cohen. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. 28(1):403–450, 2005.

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

[8] Kenneth D Harris and Alexander Thiele. Cortical state and attention. *Nature Reviews Neuroscience*, 12(9):509–523, September 2011.

[9] Seung-Hee Lee and Yang Dan. Neuromodulation of brain states. *Neuron*, 76(1):209–222, October 2012.

[10] Thomas Pfeffer, Arthur-Ervin Avramiea, Guido Nolte, Andreas K. Engel, Klaus Linkenkaer-Hansen, and Tobias H. Donner. Catecholamines alter the intrinsic variability of cortical population activity and perception. *PLOS Biology*, 16(2):e2003453, February 2018. ISSN 1545-7885. doi: 10.1371/journal.pbio.2003453. URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2003453.

[11] Matthew J McGinley, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R Cadwell, Andreas S Tolias, Jessica A Cardin, and David A McCormick. Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron*, 87(6):1143–1161, September 2015.

[12] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 1998. ISBN 978-0-262-19398-6.

[13] Leo P Sugrue, Greg S Corrado, and William T Newsome. Matching behavior and the representation of value in the parietal cortex. *Science (New York, N.Y.)*, 304(5678):1782–1787, June 2004.

[14] Samuel W. Fernberger. Interdependence of judgments within the series for the method of constant stimuli. *Journal of Experimental Psychology*, 3(2):126–150, 1920. ISSN 0022-1015(Print). doi: 10.1037/h0065212.

[15] Patrick Rabbitt and Bryan Rodgers. What does a Man do after he Makes an Error? An Analysis of Response Programming. *Quarterly Journal of Experimental Psychology*, 29(4):727–743, November 1977. ISSN 0033-555X. doi: 10.1080/14640747708400645. URL https://doi.org/10.1080/14640747708400645.

[16] Michel Treisman and Thomas C. Williams. A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, 91(1):68–111, 1984. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/0033-295X.91.1.68.

[17] Bharath Chandra Talluri, Anne E. Urai, Konstantinos Tsetsos, Marius Usher, and Tobias H. Donner. Confirmation Bias through Selective Overweighting of Choice-Consistent Evidence. *Current Biology*, 28(19):3128–3135.e8, October 2018. ISSN 0960-9822. doi: 10.1016/j.cub.2018.07.052. URL https://www.cell.com/current-biology/abstract/S0960-9822(18)30982-5.

[18] Rafal Bogacz, Eric Brown, Jeff Moehlis, Philip Holmes, and Jonathan D Cohen. 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, October 2006.

[19] Carlos D Brody and Timothy D Hanks. Neural underpinnings of the evidence accumulator. *Current opinion in neurobiology*, 37:149–157, April 2016.

[20] Michael N Shadlen and Roozbeh Kiani. Decision making as a window on cognition. *Neuron*, 80 (3):791–806, October 2013.

[21] H R Heekeren, S Marrett, and L G Ungerleider. The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, 2008.

[22] Markus Siegel, Andreas K Engel, and Tobias H Donner. Cortical network dynamics of perceptual decision-making in the human brain. *Frontiers in human neuroscience*, 5:21, 2011.

[23] Xiao-Jing Wang. Decision making in recurrent neuronal circuits. *Neuron*, 60(2):215–234, October 2008.

[24] D M Green and J A Swets. *Signal detection theory and psychophysics. 1966*. New York, 1966.

[25] Roger Ratcliff and Gail McKoon. The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, 20(4):873–922, April 2008.

[26] Tobias H Donner, Markus Siegel, Pascal Fries, and Andreas K Engel. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Current biology : CB*, 19(18): 1581–1585, September 2009.

[27] Markus Siegel, Tobias H Donner, Robert Oostenveld, Pascal Fries, and Andreas K Engel. High-frequency activity in human visual cortex is modulated by visual motion strength. *Cerebral cortex (New York, N.Y. : 1991)*, 17(3):732–741, March 2007.

[28] Valentin Wyart and Catherine Tallon-Baudry. Neural dissociation between visual awareness and spatial attention. *The Journal of Neuroscience*, 28(10):2667–2679, March 2008.

[29] Vinay Parikh, Rouba Kozak, Vicente Martinez, and Martin Sarter. Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron*, 56(1):141–154, October 2007.

[30] Peter Dayan and Angela J Yu. Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network (Bristol, England)*, 17(4):335–350, December 2006.

[31] Sebastien Bouret and Susan J Sara. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends in neurosciences*, 28(11):574–582, November 2005.

[32] Rylan S Larsen and Jack Waters. Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits*, 12, 2018.

[33] Jan Willem de Gee, Olympia Colizoli, Niels A Kloosterman, Tomas Knapen, Sander Nieuwenhuis, and Tobias H Donner. Dynamic modulation of decision biases by brainstem arousal systems. *eLife*, 6:309, April 2017.

[34] Peter R Murphy, Redmond G O'Connell, Michael O'Sullivan, Ian H Robertson, and Joshua H Balsters. Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human brain mapping*, 35(8):4140–4154, August 2014.

[35] Siddhartha Joshi, Yin Li, Rishi M Kalwani, and Joshua I Gold. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, 89(1):221–234, January 2016.

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

[37] Vincent Breton-Provencher and Mriganka Sur. Active control of arousal by a locus coeruleus GABAergic circuit. *Nature Neuroscience*, 22(2):218–228, February 2019. ISSN 1097-6256, 1546-1726. doi: 10.1038/s41593-018-0305-z. URL http://www.nature.com/articles/s41593-018-0305-z.

[38] Yang Liu, Charles Rodenkirch, Nicole Moskowitz, Brian Schriver, and Qi Wang. Dynamic Lateralization of Pupil Dilation Evoked by Locus Coeruleus Activation Results from Sympathetic, Not Parasympathetic, Contributions. *Cell reports*, 20(13):3099–3112, September 2017.

[39] Jacob Reimer, Matthew J McGinley, Yang Liu, Charles Rodenkirch, Qi Wang, David A McCormick, and Andreas S Tolias. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature communications*, 7:13289, November 2016.

[40] Jan Willem de Gee, Tomas Knapen, and Tobias H Donner. 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, February 2014.

[41] Bert Hoeks and Willem J M Levelt. Pupillary dilation as a measure of attention: a quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, 25(1):16–26, 1993.

[42] Peter R. Murphy, Evert Boonstra, and Sander Nieuwenhuis. Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nature Communications*, 7:13526, November 2016. doi: 10.1038/ncomms13526. URL https://www.nature.com/articles/ncomms13526.

[43] Samuel Cheadle, Valentin Wyart, Konstantinos Tsetsos, Nicholas Myers, Vincent de Gardelle, Santiago Herce Castañón, and Christopher Summerfield. Adaptive gain control during human perceptual choice. *Neuron*, 81(6):1429–1441, March 2014.

[44] Jan Willem de Gee, Konstantinos Tsetsos, Lars Schwabe, Anne E Urai, David A McCormick, Matthew J McGinley, and Tobias H Donner. Phasic arousal suppresses suboptimal decision biases in mice and humans. *bioRxiv*, page 447656, January 2019. doi: 10.1101/447656. URL http://biorxiv.org/content/early/2019/06/17/447656.abstract.

[45] Olympia Colizoli, Jan Willem de Gee, Anne E. Urai, and Tobias H. Donner. Task-evoked pupil responses reflect internal belief states. *Scientific Reports*, 8(1):13702, September 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-31985-3. URL https://doi.org/10.1038/s41598-018-31985-3.

[46] Anne E Urai, Jan Willem de Gee, Konstantinos Tsetsos, and Tobias H Donner. Choice history biases subsequent evidence accumulation. *eLife*, 8:e46331, July 2019. ISSN 2050-084X. doi: 10.7554/eLife.46331. URL https://doi.org/10.7554/eLife.46331.

[47] Edwin C Clayton, Janusz Rajkowski, Jonathan D Cohen, and Gary Aston-Jones. Phasic activation of monkey locus ceruleus neurons by simple decisions in a forced-choice task. *The Journal of Neuroscience*, 24(44):9914–9920, November 2004.

[48] Janusz Rajkowski, Henryk Majczynski, Edwin Clayton, and Gary Aston-Jones. Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. *Journal of neurophysiology*, 92(1):361–371, July 2004.

[49] R M Kalwani, S Joshi, and J I Gold. Phasic Activation of Individual Neurons in the Locus Ceruleus/Subceruleus Complex of Monkeys Reflects Rewarded Decisions to Go But Not Stop. *Journal of Neuroscience*, 34(41):13656–13669, October 2014.

[50] Victor de Lafuente and Ranulfo Romo. Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proceedings of the National Academy of Sciences of the United States of America*, 103(39):14266–14271, September 2006.

[51] Katharina Merten and Andreas Nieder. Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 109(16):6289–6294, April 2012.

[52] Gustavo Deco, Mar Pérez-Sanagustín, Victor de Lafuente, and Ranulfo Romo. Perceptual detection as a dynamical bistability phenomenon: A neurocomputational correlate of sensation. *Proceedings of the National Academy of Sciences*, 104(50):20073–20077, December 2007.

[53] M Usher and J L McClelland. The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, 108(3):550–592, July 2001.

[54] David Ress and David J Heeger. Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, 6(4):414–420, April 2003.

[55] Karl Zilles and Katrin Amunts. Receptor mapping: architecture of the human cerebral cortex. *Current Opinion in Neurology*, 22(4):331, August 2009. ISSN 1350-7540. doi: 10.1097/WCO.0b013e32832d95db. URL https://journals.lww.com/co-neurology/Abstract/2009/08000/Receptor_mapping__architecture_of_the_human.2.aspx.

[56] Ruud L. van den Brink, Sander Nieuwenhuis, and Tobias H. Donner. Amplification and Suppression of Distinct Brainwide Activity Patterns by Catecholamines. *Journal of Neuroscience*, 38 (34):7476–7491, August 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0514-18.2018. URL https://www.jneurosci.org/content/38/34/7476.

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

[58] Z Gil, B W Connors, and Y Amitai. Differential regulation of neocortical synapses by neuromodulators and activity. *Neuron*, 19(3):679–686, 1997.

[59] C Y Hsieh, S J Cruikshank, and R Metherate. Differential modulation of auditory thalamocortical and intracortical synaptic transmission by cholinergic agonist. *Brain research*, 880(1-2):51–64, October 2000.

[60] F Kimura, M Fukuda, and T Tsumoto. 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, October 1999.

[61] M Kobayashi, K Imamura, T Sugai, N Onoda, M Yamamoto, S Komai, and Y Watanabe. Selective suppression of horizontal propagation in rat visual cortex by norepinephrine. *The European journal of neuroscience*, 12(1):264–272, January 2000.

[62] Alexandre Pouget, Jeffrey M Beck, Wei Ji Ma, and Peter E Latham. Probabilistic brains: knowns and unknowns. *Nature Neuroscience*, 16(9):1170–1178, September 2013.

[63] Daniel J. Chandler and Barry D. Ph D. Waterhouse. Evidence for Broad Versus Segregated Projections from Cholinergic and Noradrenergic Nuclei to Functionally and Anatomically Discrete Subregions of Prefrontal Cortex. *Frontiers in Behavioral Neuroscience*, 6, 2012. ISSN 1662-5153. doi: 10.3389/fnbeh.2012.00020. URL https://www.frontiersin.org/articles/10.3389/fnbeh.2012.00020/full.

[64] Daniel J. Chandler, Wen-Jun Gao, and Barry D. Waterhouse. Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. *Proceedings of the National Academy of Sciences*, 111(18):6816–6821, May 2014. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1320827111. URL https://www.pnas.org/content/111/18/6816.

[65] Akira Uematsu, Bao Zhen Tan, Edgar A. Ycu, Jessica Sulkes Cuevas, Jenny Koivumaa, Felix Junyent, Eric J. Kremer, Ilana B. Witten, Karl Deisseroth, and Joshua P. Johansen. Modular organization of the brainstem noradrenaline system coordinates opposing learning states. *Nature Neuroscience*, 20(11):1602–1611, November 2017. ISSN 1546-1726. doi: 10.1038/nn.4642. URL https://www.nature.com/articles/nn.4642.

[66] Nelson K. Totah, Ricardo M. Neves, Stefano Panzeri, Nikos K. Logothetis, and Oxana Eschenko. The Locus Coeruleus Is a Complex and Differentiated Neuromodulatory System. *Neuron*, 99 (5):1055–1068.e6, September 2018. ISSN 0896-6273. doi: 10.1016/j.neuron.2018.07.037. URL http://www.sciencedirect.com/science/article/pii/S0896627318306354.

[67] Jeffrey M Beck, Wei Ji Ma, Xaq Pitkow, Peter E Latham, and Alexandre Pouget. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, 74(1):30–39, April 2012.

[68] Bingni W Brunton, Matthew M Botvinick, and Carlos D Brody. Rats and humans can optimally accumulate evidence for decision-making. *Science (New York, N.Y.)*, 340(6128):95–98, April 2013.

[69] A Aldo Faisal, Luc P J Selen, and Daniel M Wolpert. Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303, April 2008.

[70] I-Chun Lin, Michael Okun, Matteo Carandini, and Kenneth D Harris. The Nature of Shared Cortical Variability. *Neuron*, 87(3):644–656, August 2015.

[71] Jackson Beatty. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, 91(2):276–292, March 1982.

[72] Mark S Gilzenrat, Sander Nieuwenhuis, Marieke Jepma, and Jonathan D Cohen. 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, May 2010.

[73] Karolina M Lempert, Yu Lin Chen, and Stephen M Fleming. Relating Pupil Dilation and Metacognitive Confidence during Auditory Decision-Making. *PLOS ONE*, 10(5):e0126588, 2015.

[74] Matthew R Nassar, Katherine M Rumsey, Robert C Wilson, Kinjan Parikh, Benjamin Heasly, and Joshua I Gold. Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, 15(7):1040–1046, July 2012.

[75] Leor N Katz, Jacob L Yates, Jonathan W Pillow, and Alexander C Huk. Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, July 2016.

[76] Matthew J McGinley, Stephen V David, and David A McCormick. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron*, 87(1):179–192, July 2015.

[77] Thomas V Wiecki, Imri Sofer, and Michael J Frank. HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in neuroinformatics*, 7:14, 2013.

[78] Jacob Reimer, Emmanouil Froudarakis, Cathryn R Cadwell, Dimitri Yatsenko, George H Denfield, and Andreas S Tolias. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, 84(2):355–362, October 2014.

[79] Martin Vinck, Renata Batista-Brito, Ulf Knoblich, and Jessica A Cardin. Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron*, 86(3):740–754, May 2015.

[80] Mariana M B Cardoso, Yevgeniy B Sirotin, Bruss Lima, Elena Glushenkova, and Aniruddha Das. The neuroimaging signal is a linear sum of neurally distinct stimulus- and task-related components. *Nature Neuroscience*, 15(9):1298–1306, September 2012.

[81] Tobias H Donner, Dov Sagi, Yoram S Bonneh, and David J Heeger. Opposite neural signatures of motion-induced blindness in human dorsal and ventral visual cortex. *The Journal of Neuroscience*, 28(41):10298–10310, October 2008.

[82] D Ress, B T Backus, and D J Heeger. Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, 3(9):940–945, September 2000.

[83] Kyoung Whan Choe, Randolph Blake, and Sang-Hun Lee. Dissociation between neural signatures of stimulus and choice in population activity of human V1 during perceptual decision-making. *The Journal of Neuroscience*, 34(7):2725–2743, February 2014.

[84] Auréliane Pajani, Peter Kok, Sid Kouider, and Floris P de Lange. Spontaneous Activity Patterns in Primary Visual Cortex Predispose to Visual Hallucinations. *The Journal of Neuroscience*, 35 (37):12947–12953, September 2015.

[85] Floris P de Lange, Dobromir A Rahnev, Tobias H Donner, and Hakwan Lau. Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *The Journal of Neuroscience*, 33 (4):1400–1410, January 2013.

[86] Martin N Hebart, Tobias H Donner, and John-Dylan Haynes. Human visual and parietal cortex encode visual choices independent of motor plans. *NeuroImage*, 63(3):1393–1403, November 2012.

[87] Martin N Hebart, Yoren Schriever, Tobias H Donner, and John-Dylan Haynes. The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cerebral cortex (New York, N.Y. : 1991)*, 26(1):118–130, January 2016.

[88] Chin-An Wang, Susan E Boehnke, Brian J White, and Douglas P Munoz. Microstimulation of the monkey superior colliculus induces pupil dilation without evoking saccades. *The Journal of Neuroscience*, 32(11):3629–3636, March 2012.

[89] Laszlo Zaborszky, L Hoemke, H Mohlberg, A Schleicher, K Amunts, and K Zilles. Stereotaxic probabilistic maps of the magnocellular cell groups in human basal forebrain. *NeuroImage*, 42(3): 1127–1141, September 2008.

[90] Rodrigo A España and Craig W Berridge. Organization of noradrenergic efferents to arousal-related basal forebrain structures. *The Journal of comparative neurology*, 496(5):668–683, June 2006.

[91] Susan J Sara. The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, 10(3):211–223, March 2009.

[92] Chin-An Wang and Douglas P Munoz. A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. *Current opinion in neurobiology*, 33:134–140, August 2015.

[93] Serguei V Astafiev, Abraham Z Snyder, Gordon L Shulman, and Maurizio Corbetta. Comment on "Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI" and "Homeostatic sleep pressure and responses to sustained attention in the suprachiasmatic area". *Science (New York, N.Y.)*, 328(5976):309–author reply 309, April 2010.

[94] F Beissner. Functional MRI of the Brainstem: Common Problems and their Solutions. *Clinical neuroradiology*, 25 Suppl 2:251–257, October 2015.

[95] Jonathan C W Brooks, Olivia K Faull, Kyle T S Pattinson, and Mark Jenkinson. Physiological noise in brainstem FMRI. *Frontiers in human neuroscience*, 7:623, 2013.

[96] Birte U Forstmann, Gilles de Hollander, Leendert van Maanen, Anneke Alkemade, and Max C Keuken. Towards a mechanistic understanding of the human subcortex. *Nature Reviews Neuroscience*, 18(1):57–65, January 2017.

[97] Noam I Keren, Carl T Lozar, Kelly C Harris, Paul S Morgan, and Mark A Eckert. In vivo mapping of the human locus coeruleus. *NeuroImage*, 47(4):1261–1267, October 2009.

[98] Noam I Keren, Saeid Taheri, Elena M Vazey, Paul S Morgan, Ann-Charlotte E Granholm, Gary S Aston-Jones, and Mark A Eckert. Histologic validation of locus coeruleus MRI contrast in post-mortem tissue. *NeuroImage*, 113(C):235–245, June 2015.

[99] M. A. Eckert, N. I. Keren, and G. Aston-Jones. Looking forward with the locus coeruleus. *Science E-Letter*, 2010.

[100] Christopher M Warren, Eran Eldar, Ruud L van den Brink, Klodianna-Daphne Tona, Nic J van der Wee, Eric J Giltay, Martijn S van Noorden, Jos A Bosch, Robert C Wilson, Jonathan D Cohen, and Sander Nieuwenhuis. Catecholamine-Mediated Increases in Gain Enhance the Precision of Cortical Representations. *The Journal of Neuroscience*, 36(21):5699–5708, May 2016.

[101] Peter R Murphy, Joachim Vandekerckhove, and Sander Nieuwenhuis. Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. *PLoS computational biology*, 10(9):e1003854, September 2014.

[102] Christopher D Fiorillo, Philippe N Tobler, and Wolfram Schultz. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, N.Y.)*, 299(5614):1898–1902, March 2003.

[103] Kensaku Nomoto, Wolfram Schultz, Takeo Watanabe, and Masamichi Sakagami. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *The Journal of Neuroscience*, 30(32):10692–10702, August 2010.

[104] Eran Eldar, Jonathan D Cohen, and Yael Niv. The effects of neural gain on attention and learning. *Nature Neuroscience*, 16(8):1146–1153, August 2013.

[105] Rosalyn J Moran, Pablo Campo, Mkael Symmonds, Klaas E Stephan, Raymond J Dolan, and Karl J Friston. Free energy, precision and learning: the role of cholinergic neuromodulation. *The Journal of Neuroscience*, 33(19):8227–8236, May 2013.

[106] Praveen K Pilly and Aaron R Seitz. What a difference a parameter makes: a psychophysical comparison of random dot motion algorithms. *Vision research*, 49(13):1599–1612, June 2009.

[107] Eri Shibata, Makoto Sasaki, Koujiro Tohyama, Kotaro Otsuka, and Akio Sakai. Reduced signal of locus ceruleus in depression in quantitative neuromelanin magnetic resonance imaging. *Neuroreport*, 18(5):415–418, March 2007.

[108] Tomas Knapen, Jan Willem de Gee, Jan Brascamp, Stijn Nuiten, Sylco Hoppenbrouwers, and Jan Theeuwes. Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance. *PLOS ONE*, 11(5):e0155574, 2016.

[109] Christoph W Korn and Dominik R Bach. A solid frame for the window on cognition: Modeling event-related pupil responses. *Journal of vision*, 16(3):28, 2016.

[110] D H McDougal and PDR Gamlin. *Pupillary control pathways*. The senses: A comprehensive reference, 2008.

[111] Norman R Draper and Harry Smith. *Applied Regression Analysis*. Wiley-Interscience, April 1998.

[112] Roger Ratcliff and Russ Childers. Individual Differences and Fitting Methods for the Two-Choice Diffusion Model of Decision Making. *Decision (Washington, D.C.)*, 2015, 2015.

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

[114] Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy E J Behrens, Heidi Johansen-Berg, Peter R Bannister, Marilena De Luca, Ivana Drobnjak, David E Flitney, Rami K Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J Michael Brady, and Paul M Matthews. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1:S208–19, 2004.

[115] A M Dale, B Fischl, and M I Sereno. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, February 1999.

[116] B Fischl, M I Sereno, and A M Dale. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, February 1999.

[117] M Jenkinson, P Bannister, M Brady, and S Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 2002.

[118] G H Glover, T Q Li, and D Ress. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic resonance in medicine*, 2000.

[119] C C Chang and C J Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and \ldots*, 2011.

[120] Anders Eklund, Thomas E Nichols, and Hans Knutsson. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, 113(28):7900–7905, July 2016.

[121] Vishnu P Murty, Maheen Shermohammed, David V Smith, R McKell Carter, Scott A Huettel, and R Alison Adcock. Resting state networks distinguish human ventral tegmental area from substantia nigra. *NeuroImage*, 100:580–589, October 2014.

[122] Simon B Eickhoff, Klaas E Stephan, Hartmut Mohlberg, Christian Grefkes, Gereon R Fink, Katrin Amunts, and Karl Zilles. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4):1325–1335, May 2005.

[123] Christophe Destrieux, Bruce Fischl, Anders Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, 53(1):1–15, October 2010.

[124] Serge O Dumoulin and Brian A Wandell. Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2):647–660, January 2008.

[125] Anne E Urai, Anke Braun, and Tobias H Donner. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature communications*, 8:14637, March 2017.

[126] Kamesh Krishnamurthy, Matthew R Nassar, Shilpa Sarode, and Joshua I Gold. Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nature Human Behaviour*, 1:0107, 2017.

[127] Armin Lak, Kensaku Nomoto, Mehdi Keramati, Masamichi Sakagami, and Adam Kepecs. Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. *Current biology : CB*, 27(6):821–832, March 2017.

[128] Daniel Kahneman. *Thinking, Fast and Slow*. Macmillan, October 2011.

[129] D Amaral and H Sinnamon. The locus coeruleus: neurobiology of a central noradrenergic nucleus. *Progress in Neurobiology*, 9(3):147–196, 1977. ISSN 03010082. doi: 10.1016/0301-0082(77)90016-8. URL https://linkinghub.elsevier.com/retrieve/pii/0301008277900168.

[130] Craig W Berridge and Barry D Waterhouse. The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain research. Brain research reviews*, 42(1):33–84, April 2003.

[131] Michael N. Shadlen and Daphna Shohamy. Decision Making and Sequential Sampling from Memory. *Neuron*, 90(5):927–939, June 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.04.036. URL http://www.sciencedirect.com/science/article/pii/S0896627316301234.

[132] Konstantinos Tsetsos, Nick Chater, and Marius Usher. 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, June 2012.

[133] Jean-Michel Hupé, Cédric Lamirel, and Jean Lorenceau. Pupil dynamics during bistable motion perception. *Journal of vision*, 9(7):10, July 2009.

[134] Christian R Lee and David J Margolis. Pupil Dynamics Reflect Behavioral Choice and Learning in a Go/NoGo Tactile Decision-Making Task in Mice. *Frontiers in behavioral neuroscience*, 10:200, 2016.

[135] Brian J. Schriver, Svetlana Bagdasarov, and Qi Wang. Pupil-linked arousal modulates behavior in rats performing a whisker deflection direction discrimination task. *Journal of Neurophysiology*, 120(4):1655–1670, October 2018. ISSN 1522-1598. doi: 10.1152/jn.00290.2018.

[136] Robert M Yerkes and John D Dodson. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5):459–482, November 1908.

[137] Roger Ratcliff, Cynthia Huang-Pollock, and Gail McKoon. Modeling Individual Differences in the Go/No-Go Task With a Diffusion Model. *Decision*, 2016.

[138] Jan Drugowitsch, Valentin Wyart, Anne-Dominique Devauchelle, and Etienne Koechlin. Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, 92(6):1398–1411, December 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2016.11.005. URL http://www.sciencedirect.com/science/article/pii/S0896627316308431.

[139] Roger Ratcliff. A theory of memory retrieval. *Psychological Review*, 85(2):59–108, 1978. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/0033-295X.85.2.59.

[140] Holly J. Bowen, Julia Spaniol, Ronak Patel, and Andreas Voss. A Diffusion Model Analysis of Decision Biases Affecting Delayed Recognition of Emotional Stimuli. *PLOS ONE*, 11(1):e0146769, January 2016. ISSN 1932-6203. doi: 10.1371/journal.pone.0146769. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146769.

[141] Konstantinos Tsetsos, Rani Moran, James Moreland, Nick Chater, Marius Usher, and Christopher Summerfield. 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, March 2016.

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

[143] Gideon Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, March 1978.

[144] Kong-Fatt Wong and Xiao-Jing Wang. A recurrent network mechanism of time integration in perceptual decisions. *The Journal of Neuroscience*, 26(4):1314–1328, January 2006.

[145] Hagar Gelbard-Sagiv, Efrat Magidov, Haggai Sharon, Talma Hendler, and Yuval Nir. Noradrenaline Modulates Visual Perception and Late Visually Evoked Activity. *Current Biology*, 28 (14):2239–2249.e6, July 2018. ISSN 09609822. doi: 10.1016/j.cub.2018.05.051. URL https://linkinghub.elsevier.com/retrieve/pii/S0960982218306870.

[146] Dougal G.R. Tervo, Mikhail Proskurin, Maxim Manakov, Mayank Kabra, Alison Vollmer, Kristin Branson, and Alla Y. Karpova. Behavioral Variability through Stochastic Choice and Its Gating by Anterior Cingulate Cortex. *Cell*, 159(1):21–32, September 2014. ISSN 00928674. doi: 10.1016/j.cell.2014.08.037. URL https://linkinghub.elsevier.com/retrieve/pii/S0092867414011076.

[147] Gary A. Kane, Elena M. Vazey, Robert C. Wilson, Amitai Shenhav, Nathaniel D. Daw, Gary Aston-Jones, and Jonathan D. Cohen. Increased locus coeruleus tonic activity causes disengagement from a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, 17(6):1073–1083, December 2017. ISSN 1530-7026, 1531-135X. doi: 10.3758/s13415-017-0531-y. URL http://link.springer.com/10.3758/s13415-017-0531-y.

[148] Caroline I. Jahn, Sophie Gilardeau, Chiara Varazzani, Bastien Blain, Jerome Sallet, Mark E. Walton, and Sebastien Bouret. Dual contributions of noradrenaline to behavioural flexibility and motivation. *Psychopharmacology*, 235(9):2687–2702, September 2018. ISSN 0033-3158, 1432-2072. doi: 10.1007/s00213-018-4963-z. URL http://link.springer.com/10.1007/s00213-018-4963-z.

[149] Michael E. Hasselmo. The Role of Acetylcholine in Learning and Memory. *Current opinion in neurobiology*, 16(6):710–715, December 2006. ISSN 0959-4388. doi: 10.1016/j.conb.2006.09.002. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2659740/.

[150] Anthony J Porcelli and Mauricio R Delgado. Stress and decision making: effects on valuation, learning, and risk-taking. *Current Opinion in Behavioral Sciences*, 14:33–39, April 2017. ISSN 23521546. doi: 10.1016/j.cobeha.2016.11.015. URL https://linkinghub.elsevier.com/retrieve/pii/S2352154616301905.

[151] Anthony J. Porcelli and Mauricio R. Delgado. Acute Stress Modulates Risk Taking in Financial Decision Making. *Psychological Science*, 20(3):278–283, March 2009. ISSN 0956-7976, 1467-9280. doi: 10.1111/j.1467-9280.2009.02288.x. URL http://journals.sagepub.com/doi/10.1111/j.1467-9280.2009.02288.x.

[152] Jennifer K. Lenow, Sara M. Constantino, Nathaniel D. Daw, and Elizabeth A. Phelps. Chronic and Acute Stress Promote Overexploitation in Serial Decision Making. *The Journal of Neuroscience*, 37(23):5681–5689, June 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3618-16.2017. URL http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3618-16.2017.

[153] A. R. Otto, C. M. Raio, A. Chiang, E. A. Phelps, and N. D. Daw. Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, 110

(52):20941–20946, December 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1312011110. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1312011110.

[154] L. Schwabe, O. Hoffken, M. Tegenthoff, and O. T. Wolf. Preventing the Stress-Induced Shift from Goal-Directed to Habit Action with a -Adrenergic Antagonist. *Journal of Neuroscience*, 31(47): 17317–17325, November 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3304-11.2011. URL http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3304-11.2011.

[155] Amy F. T. Arnsten. Stress weakens prefrontal networks: molecular insults to higher cognition. *Nature Neuroscience*, 18(10):1376–1385, October 2015. ISSN 1546-1726. doi: 10.1038/nn.4087. URL https://www.nature.com/articles/nn.4087.

[156] Adam Kepecs, Naoshige Uchida, Hatim A Zariwala, and Zachary F Mainen. Neural correlates, computation and behavioural impact of decision confidence. *Nature*, 455(7210):227–231, September 2008.

[157] Wei Ji Ma and Mehrdad Jazayeri. Neural coding of uncertainty and probability. *Annual review of neuroscience*, 37:205–220, 2014.

[158] Alexandre Pouget, Jan Drugowitsch, and Adam Kepecs. Confidence and certainty: distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3):366–374, March 2016. ISSN 1097-6256, 1546-1726. doi: 10.1038/nn.4240. URL http://www.nature.com/articles/nn.4240.

[159] Anne Bergt, Anne E. Urai, Tobias H. Donner, and Lars Schwabe. Reading memory formation from the eyes. *European Journal of Neuroscience*, 47(12):1525–1533, June 2018. ISSN 0953816X. doi: 10.1111/ejn.13984. URL http://doi.wiley.com/10.1111/ejn.13984.

[160] Kentaro Katahira. How hierarchical models improve point estimates of model parameters at the individual level. *Journal of Mathematical Psychology*, 73:37–58, August 2016. ISSN 0022-2496. doi: 10.1016/j.jmp.2016.03.007. URL http://www.sciencedirect.com/science/article/pii/S0022249616300025.

[161] Joachim Vandekerckhove, Francis Tuerlinckx, and Michael D. Lee. Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 16(1):44–62, 2011. ISSN 1939-1463, 1082-989X. doi: 10.1037/a0021765. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0021765.

[162] Moshe Glickman, Konstantinos Tsetsos, and Marius Usher. Attentional Selection Mediates Framing and Risk-Bias Effects. *Psychological Science*, 29(12):2010–2019, December 2018. ISSN 0956-7976, 1467-9280. doi: 10.1177/0956797618803643. URL http://journals.sagepub.com/doi/10.1177/0956797618803643.

[163] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting Linear Mixed-Effects Models Usinglme4. *Journal of Statistical Software*, 67(1), 2015.

[164] Dale J Barr, Roger Levy, Christoph Scheepers, and Harry J Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3):255–278, April 2013.

[165] Alexandra Kuznetsova, Per B Brockhoff, and Rune H B Christensen. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 2017.

[166] Miaomiao Jin and Lindsey L. Glickfeld. Contribution of Sensory Encoding to Measured Bias. *Journal of Neuroscience*, 39(26):5115–5127, June 2019. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0076-19.2019. URL http://www.jneurosci.org/content/39/26/5115.

[167] Tolgay Ergenoglu, Tamer Demiralp, Zubeyir Bayraktaroglu, Mehmet Ergen, Huseyin Beydagi, and Yagiz Uresin. Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research*, 20(3):376–383, August 2004. ISSN 0926-6410. doi: 10.1016/j.cogbrainres.2004.03.009. URL http://www.sciencedirect.com/science/article/pii/S0926641004000941.

[168] Luca Iemi, Maximilien Chaumon, Sébastien M. Crouzet, and Niko A. Busch. Spontaneous Neural Oscillations Bias Perception by Modulating Baseline Excitability. *Journal of Neuroscience*, 37(4):807–819, January 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1432-16.2016. URL http://www.jneurosci.org/content/37/4/807.

[169] Niels A Kloosterman, Jan Willem de Gee, Markus Werkle-Bergner, Ulman Lindenberger, Douglas D Garrett, and Johannes Jacobus Fahrenfort. Humans strategically shift decision bias by flexibly adjusting sensory evidence accumulation. *eLife*, 8:e37321, February 2019. ISSN 2050-084X. doi: 10.7554/eLife.37321. URL https://doi.org/10.7554/eLife.37321.

[170] Katharina Limbach and Paul M. Corballis. Prestimulus alpha power influences response criterion in a detection task. *Psychophysiology*, 53(8):1154–1164, 2016. ISSN 1469-8986. doi: 10.1111/psyp.12666. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/psyp.12666.

[171] Thomas Meindertsma, Niels A Kloosterman, Guido Nolte, Andreas K Engel, and Tobias H Donner. Multiple Transient Signals in Human Visual Cortex Associated with an Elementary Decision. *The Journal of Neuroscience*, 37(23):5744–5757, June 2017.

[172] A B Watson and D G Pelli. QUEST: A Bayesian adaptive psychometric method. *Perception & psychophysics*, 1983.

[173] Koryna Lewandowska, Adam Gągol, Barbara Sikora-Wachowicz, Tadeusz Marek, and Magdalena Fąfrowicz. Saying "yes" when you want to say "no" - pupil dilation reflects evidence accumulation in a visual working memory recognition task. *International Journal of Psychophysiology*, 139:18–32, May 2019. ISSN 01678760. doi: 10.1016/j.ijpsycho.2019.03.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0167876018308547.

[174] Iain Stitt, Zhe Charles Zhou, Susanne Radtke-Schuller, and Flavio Fröhlich. Arousal dependent modulation of thalamo-cortical functional interaction. *Nature Communications*, 9(1):2455, June 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-04785-6. URL https://www.nature.com/articles/s41467-018-04785-6.

[175] Peter R Murphy, Marianne L van Moort, and Sander Nieuwenhuis. The Pupillary Orienting Response Predicts Adaptive Behavioral Adjustment after Errors. *PLOS ONE*, 11(3):e0151763, 2016.

[176] Joachim Lange, Robert Oostenveld, and Pascal Fries. Reduced Occipital Alpha Power Indexes Enhanced Excitability Rather than Improved Visual Perception. *Journal of Neuroscience*, 33(7): 3212–3220, February 2013. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3755-12.2013. URL http://www.jneurosci.org/content/33/7/3212.

[177] Georgios Michalareas, Julien Vezoli, Stan van Pelt, Jan-Mathijs Schoffelen, Henry Kennedy, and Pascal Fries. Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron*, 89(2):384–397, January 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.12.018. URL http://www.sciencedirect.com/science/article/pii/S0896627315011204.

[178] Markus Siegel, Tobias H Donner, and Andreas K Engel. Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2):121–134, February 2012.

[179] Timo van Kerkoerle, Matthew W. Self, Bruno Dagnino, Marie-Alice Gariel-Mathis, Jasper Poort, Chris van der Togt, and Pieter R. Roelfsema. Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, 111(40):14332–14341, October 2014. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1402773111. URL http://www.pnas.org/lookup/doi/10.1073/pnas.1402773111.

[180] Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Intell. Neuroscience*, 2011:1:1–1:9, January 2011. ISSN 1687-5265. doi: 10.1155/2011/156869. URL http://dx.doi.org/10.1155/2011/156869.

[181] A M Dale. Optimal experimental design for event-related fMRI. *Human brain mapping*, 8(2-3): 109–114, 1999.

[182] Matthew F. Glasser, Timothy S. Coalson, Emma C. Robinson, Carl D. Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F. Beckmann, Mark Jenkinson, Stephen M. Smith, and David C. Van Essen. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, August 2016. ISSN 1476-4687. doi: 10.1038/nature18933. URL https://www.nature.com/articles/nature18933.

[183] Liang Wang, Ryan E B Mruczek, Michael J Arcaro, and Sabine Kastner. Probabilistic Maps of Visual Topography in Human Cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 25(10):3911–3931, September 2015.

[184] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Lauri Parkkonen, and Matti S Hämäläinen. MNE software for processing MEG and EEG data. *NeuroImage*, 86:446–460, February 2014.

[185] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44(9):867–880, September 1997. ISSN 0018-9294. doi: 10.1109/10.623056.

[186] P. P. Mitra and B. Pesaran. Analysis of Dynamic Brain Imaging Data. *Biophysical Journal*, 76 (2):691–708, February 1999. ISSN 0006-3495. doi: 10.1016/S0006-3495(99)77236-X. URL http://www.sciencedirect.com/science/article/pii/S000634959977236X.

[187] Brian A Wandell, Serge O Dumoulin, and Alyssa A Brewer. Visual Field Maps in Human Cortex. *Neuron*, 56(2):366–383, December 2006.

[188] Y Rosseel. lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 2012.

[189] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55, April 2016. ISSN 2376-5992. doi: 10.7717/peerj-cs.55. URL https://peerj.com/articles/cs-55.

[190] Tal Yarkoni and Jake Westfall. Bambi: A simple interface for fitting Bayesian mixed effects models. September 2016. doi: 10.31219/osf.io/rv7sn. URL https://osf.io/rv7sn/.

[191] Eric Maris and Robert Oostenveld. Nonparametric statistical testing of EEG- and MEG-data. *Journal of neuroscience methods*, 164(1):177–190, August 2007.

[192] Florent Meyniel, Mariano Sigman, and Zachary F Mainen. Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, 88(1):78–92, October 2015.

[193] P. Read Montague, Steven E. Hyman, and Jonathan D. Cohen. Computational roles for dopamine in behavioural control. *Nature*, 431(7010):760–767, October 2004. ISSN 0028-0836, 1476-4687. doi: 10.1038/nature03015. URL http://www.nature.com/articles/nature03015.

[194] Wolfram Schultz. Behavioral Theories and the Neurophysiology of Reward. *Annual Review of Psychology*, 57(1):87–115, January 2006. ISSN 0066-4308, 1545-2085. doi: 10.1146/annurev.psych.56.091103.070229. URL http://www.annualreviews.org/doi/10.1146/annurev.psych.56.091103.070229.

[195] P. W. Glimcher. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(Supplement_3):15647–15654, September 2011. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1014269108. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1014269108.

[196] Wolfram Schultz. Getting Formal with Dopamine and Reward. *Neuron*, 36(2):241–263, October 2002. ISSN 08966273. doi: 10.1016/S0896-6273(02)00967-4. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627302009674.

[197] Angela J Yu and Peter Dayan. Uncertainty, neuromodulation, and attention. *Neuron*, 46(4):681–692, May 2005.

[198] P. Dayan and N. D. Daw. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):429–453, December 2008. ISSN 1530-7026, 1531-135X. doi: 10.3758/CABN.8.4.429. URL http://www.springerlink.com/index/10.3758/CABN.8.4.429.

[199] K Preuschoff and B Marius't Hart. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Frontiers in \ldots*, 2011.

[200] J. X. O'Reilly, U. Schuffelgen, S. F. Cuell, T. E. J. Behrens, R. B. Mars, and M. F. S. Rushworth. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 110(38):E3660–E3669, September 2013. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1305373110. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1305373110.

[201] Jackson Beatty and Daniel Kahneman. Pupillary changes in two memory tasks. *Psychonomic Science*, 5(10):371–372, 1966.

[202] H. M. Simpson and Shirley M. Hale. Pupillary Changes during a Decision-Making Task. *Perceptual and Motor Skills*, 29(2):495–498, October 1969. ISSN 0031-5125. doi: 10.2466/pms.1969.29.2.495. URL https://doi.org/10.2466/pms.1969.29.2.495.

[203] GAD Hakerem and Samuel Sutton. Pupillary Response at Visual Threshold. *Nature*, 212(5061):485–486, October 1966.

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

[205] Eckhard H. Hess and James M. Polt. Pupil Size as Related to Interest Value of Visual Stimuli. *Science*, 132(3423):349–350, August 1960. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.132.3423.349. URL https://science.sciencemag.org/content/132/3423/349.

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

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

[208] Anke Braun, Anne E. Urai, and Tobias H. Donner. Adaptive History Biases Result from Confidence-Weighted Accumulation of past Choices. *Journal of Neuroscience*, 38(10):2418–2429, March 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2189-17.2017. URL http://www.jneurosci.org/content/38/10/2418.

[209] Joshua I. Sanders, Balázs Hangya, and Adam Kepecs. Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron*, 90(3):499–506, May 2016. ISSN 08966273. doi: 10.1016/j.neuron.2016.03.025. URL https://linkinghub.elsevier.com/retrieve/pii/S0896627316300162.

[210] Anne E. Urai and Klaus Wimmer. Spatiotemporal motion energy filtering: a Matlab implementation, November 2016. URL https://zenodo.org/record/167351#.XLuYPJNKh24.

[211] Archy O. de Berker, Robb B. Rutledge, Christoph Mathys, Louise Marshall, Gemma F. Cross, Raymond J. Dolan, and Sven Bestmann. Computations of uncertainty mediate acute stress responses in humans. *Nature Communications*, 7(1):10996, December 2016. ISSN 2041-1723. doi: 10.1038/ncomms10996. URL http://www.nature.com/articles/ncomms10996.

[212] Jan R. Wessel, Claudia Danielmeier, and Markus Ullsperger. Error Awareness Revisited: Accumulation of Multimodal Evidence from Central and Autonomic Nervous Systems. *Journal of Cognitive Neuroscience*, 23(10):3021–3036, October 2011. ISSN 0898-929X, 1530-8898. doi: 10.1162/jocn.2011.21635. URL http://www.mitpressjournals.org/doi/10.1162/jocn.2011.21635.

[213] Joanne C. Van Slooten, Sara Jahfari, Tomas Knapen, and Jan Theeuwes. How pupil responses track value-based decision-making during and after reinforcement learning. *PLOS Computational Biology*, 14(11):e1006632, November 2018. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1006632. URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006632.

[214] Mark F. Bear and Wolf Singer. Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature*, 320(6058):172, March 1986. ISSN 1476-4687. doi: 10.1038/320172a0. URL https://www.nature.com/articles/320172a0.

[215] Pieter R Roelfsema and Anthony Holtmaat. Control of synaptic plasticity in deep cortical networks. *Nature Reviews Neuroscience*, 19(3):166–180, March 2018.

[216] David Weinshenker and Jason P Schroeder. There and back again: a tale of norepinephrine and drug addiction. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, 32(7):1433–1451, July 2007.

[217] Sebastiaan Mathot. Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, 1(1):16, February 2018. ISSN 2514-4820. doi: 10.5334/joc.18. URL http://www.journalofcognition.org/articles/10.5334/joc.18/.

[218] S. G. Manohar and M. Husain. Reduced pupillary reward sensitivity in Parkinson's disease. *npj Parkinson's Disease*, 1:15026, December 2015. ISSN 2373-8057. doi: 10.1038/npjparkd.2015.26. URL https://www.nature.com/articles/npjparkd201526.

[219] Wolfram Schultz. Multiple Dopamine Functions at Different Time Courses. *Annual Review of Neuroscience*, 30(1):259–288, July 2007. ISSN 0147-006X, 1545-4126. doi: 10.1146/annurev.neuro.28.061604.135722. URL http://www.annualreviews.org/doi/10.1146/annurev.neuro.28.061604.135722.

[220] Chi-Tat Law and Joshua I. Gold. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, 12(5):655–663, May 2009. ISSN 1546-1726. doi: 10.1038/nn.2304. URL https://www.nature.com/articles/nn.2304.

[221] Pieter R Roelfsema, Arjen van Ooyen, and Takeo Watanabe. Perceptual learning rules based on reinforcers and attention. *Trends in cognitive sciences*, 14(2):64–71, February 2010.

[222] Stephen M. Fleming and Hakwan C. Lau. How to measure metacognition. *Frontiers in Human Neuroscience*, 8, 2014. ISSN 1662-5161. doi: 10.3389/fnhum.2014.00443. URL https://www.frontiersin.org/articles/10.3389/fnhum.2014.00443/full.

[223] Ariel Zylberberg, Pieter R Roelfsema, and Mariano Sigman. Variance misperception explains illusions of confidence in simple perceptual decisions. *Consciousness and cognition*, 27:246–253, July 2014.

[224] Maël Lebreton, Shari Langdon, Matthijs J Slieker, Jip S Nooitgedacht, Anna E Goudriaan, Damiaan Denys, Ruth J van Holst, and Judy Luigjes. Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *SCIENCE ADVANCES*, page 14, 2018.

[225] Gilles Dutilh, Don van Ravenzwaaij, Sander Nieuwenhuis, Han L. J. van der Maas, Birte U. Forstmann, and Eric-Jan Wagenmakers. How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology*, 56(3):208–216, June 2012. ISSN 0022-2496. doi: 10.1016/j.jmp.2012.04.001. URL http://www.sciencedirect.com/science/article/pii/S0022249612000454.

[226] Kobe Desender, Peter Murphy, Annika Boldt, Tom Verguts, and Nick Yeung. A Postdecisional Neural Marker of Confidence Predicts Information-Seeking in Decision-Making. *Journal of Neuroscience*, 39(17):3309–3319, April 2019. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.2620-18.2019. URL http://www.jneurosci.org/content/39/17/3309.

[227] Rei Akaishi, Kazumasa Umeda, Asako Nagase, and Katsuyuki Sakai. Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia. *Neuron*, 81(1):195–206, January 2014. ISSN 0896-6273. doi: 10.1016/j.neuron.2013.10.018. URL http://www.sciencedirect.com/science/article/pii/S0896627313009215.

[228] Ingo Fründ, Felix A. Wichmann, and Jakob H. Macke. Quantifying the effect of intertrial dependence on perceptual decisions. *Journal of Vision*, 14(7), June 2014. ISSN 1534-7362. doi: 10.1167/14.7.9.

[229] J. I. Gold, C. T. Law, P. Connolly, and S. Bennur. The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *Journal of neurophysiology*, 100(5):2653–2668, November 2008. ISSN 0022-3077. doi: 10.1152/jn.90629.2008. URL https://europepmc.org/articles/PMC2585410/.

[230] Laura Busse, Asli Ayaz, Neel T. Dhruv, Steffen Katzner, Aman B. Saleem, Marieke L. Schölvinck, Andrew D. Zaharia, and Matteo Carandini. The Detection of Visual Contrast in the Behaving Mouse. *Journal of Neuroscience*, 31(31):11351–11361, August 2011. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.6689-10.2011. URL http://www.jneurosci.org/content/31/31/11351.

[231] Onyekachi Odoemene, Sashank Pisupati, Hien Nguyen, and Anne K. Churchland. Visual Evidence Accumulation Guides Decision-Making in Unrestrained Mice. *The Journal of Neuroscience*, 38(47):10143–10155, November 2018. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3478-17.2018. URL http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3478-17.2018.

[232] Arman Abrahamyan, Laura Luz Silva, Steven C. Dakin, Matteo Carandini, and Justin L. Gardner. Adaptable history biases in human perceptual decisions. *Proceedings of the National Academy*

of Sciences, 113(25):E3548–E3557, June 2016. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1518786113. URL https://www.pnas.org/content/113/25/E3548.

[233] Anna-Antonia Pape and Markus Siegel. Motor cortex activity predicts response alternation during sensorimotor decisions. *Nature Communications*, 7:13098, October 2016. ISSN 2041-1723. doi: 10.1038/ncomms13098. URL https://www.nature.com/articles/ncomms13098.

[234] Braden A. Purcell and Roozbeh Kiani. Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex. *Neuron*, 89(3):658–671, February 2016. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.12.027. URL http://www.sciencedirect.com/science/article/pii/S0896627315011290.

[235] Elexa St John-Saaltink, Peter Kok, Hakwan C. Lau, and Floris P. de Lange. Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *Journal of Neuroscience*, 36(23):6186–6192, June 2016. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.4390-15.2016. URL http://www.jneurosci.org/content/36/23/6186.

[236] D. Thura, G. Guberman, and P. Cisek. Trial-to-trial adjustments of speed-accuracy trade-offs in premotor and primary motor cortex. *Journal of neurophysiology*, 117(2):665–683, February 2017. ISSN 0022-3077. doi: 10.1152/jn.00726.2016. URL https://europepmc.org/articles/PMC5288471/.

[237] Eun Jung Hwang, Jeffrey E. Dahlen, Madan Mukundan, and Takaki Komiyama. History-based action selection bias in posterior parietal cortex. *Nature Communications*, 8(1):1242, November 2017. ISSN 2041-1723. doi: 10.1038/s41467-017-01356-z. URL https://www.nature.com/articles/s41467-017-01356-z.

[238] Benjamin B. Scott, Christine M. Constantinople, Athena Akrami, Timothy D. Hanks, Carlos D. Brody, and David W. Tank. Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. *Neuron*, 95(2):385–398.e5, July 2017. ISSN 0896-6273. doi: 10.1016/j.neuron.2017.06.013. URL http://www.sciencedirect.com/science/article/pii/S0896627317305111.

[239] Angela J Yu and Jonathan D Cohen. Sequential effects: Superstition or rational behavior? In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 1873–1880. Curran Associates, Inc., 2009. URL http://papers.nips.cc/paper/3519-sequential-effects-superstition-or-rational-behavior.pdf.

[240] Shunan Zhang, He Crane Huang, and Angela J Yu. Sequential effects: A Bayesian analysis of prior bias on reaction time and behavioral choice. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36:7, 2014.

[241] Christopher M Glaze, Joseph W Kable, and Joshua I Gold. Normative evidence accumulation in unpredictable environments. *eLife*, 4:e08825, 2015.

[242] Raymond Y. Cho, Leigh E. Nystrom, Eric T. Brown, Andrew D. Jones, Todd S. Braver, Philip J. Holmes, and Jonathan D. Cohen. Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. *Cognitive, Affective, & Behavioral Neuro-*

science, 2(4):283–299, December 2002. ISSN 1531-135X. doi: 10.3758/CABN.2.4.283. URL https://doi.org/10.3758/CABN.2.4.283.

[243] Juan Gao, KongFatt Wong-Lin, Philip Holmes, Patrick Simen, and Jonathan D. Cohen. Sequential Effects in Two-Choice Reaction Time Tasks: Decomposition and Synthesis of Mechanisms. *Neural Computation*, 21(9):2407–2436, June 2009. ISSN 0899-7667. doi: 10.1162/neco.2009.09-08-866. URL https://doi.org/10.1162/neco.2009.09-08-866.

[244] Juan Gao, Rebecca Tortell, and James L. McClelland. Dynamic Integration of Reward and Stimulus Information in Perceptual Decision-Making. *PLoS ONE*, 6(3), March 2011. ISSN 1932-6203. doi: 10.1371/journal.pone.0016749. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3048391/.

[245] James J Bonaiuto, Archy de Berker, and Sven Bestmann. Response repetition biases in human perceptual decisions are explained by activity decay in competitive attractor models. *eLife*, 5:e20047, December 2016. ISSN 2050-084X. doi: 10.7554/eLife.20047. URL https://elifesciences.org/articles/20047.

[246] Matthew Wilder, Matt Jones, and Michael C Mozer. Sequential effects reflect parallel learning of multiple environmental regularities. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems 22*, pages 2053–2061. Curran Associates, Inc., 2009. URL http://papers.nips.cc/paper/3870-sequential-effects-reflect-parallel-learning-of-multiple-environmental-regularit pdf.

[247] Stefan Bode, David K. Sewell, Simon Lilburn, Jason D. Forte, Philip L. Smith, and Jutta Stahl. Predicting Perceptual Decision Biases from Early Brain Activity. *Journal of Neuroscience*, 32(36):12488–12498, September 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.1708-12.2012. URL http://www.jneurosci.org/content/32/36/12488.

[248] Matt Jones, Tim Curran, Michael C. Mozer, and Matthew H. Wilder. Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review*, 120(3):628–666, 2013. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/a0033180.

[249] Corey N. White and Russell A. Poldrack. Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 40(2):385–398, March 2014. ISSN 1939-1285. doi: 10.1037/a0034851.

[250] Roger Ratcliff. Modeling response signal and response time data. *Cognitive Psychology*, 53(3):195–237, November 2006. ISSN 0010-0285. doi: 10.1016/j.cogpsych.2005.10.002. URL http://www.sciencedirect.com/science/article/pii/S0010028505000757.

[251] Sara Jahfari, Frederick Verbruggen, Michael J. Frank, Lourens J. Waldorp, Lorenza Colzato, K. Richard Ridderinkhof, and Birte U. Forstmann. How Preparation Changes the Need for Top–Down Control of the Basal Ganglia When Inhibiting Premature Actions. *Journal of Neuroscience*, 32(32):10870–10878, August 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0902-12.2012. URL http://www.jneurosci.org/content/32/32/10870.

[252] Roozbeh Kiani, Timothy D. Hanks, and Michael N. Shadlen. Bounded Integration in Parietal Cortex Underlies Decisions Even When Viewing Duration Is Dictated by the Environment. *Journal of Neuroscience*, 28(12):3017–3029, March 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.4761-07.2008. URL http://www.jneurosci.org/content/28/12/3017.

[253] Konstantinos Tsetsos, Thomas Pfeffer, Pia Jentgens, and Tobias H. Donner. Action Planning and the Timescale of Evidence Accumulation. *PLoS ONE*, 10(6), June 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0129473. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4467085/.

[254] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, December 1974. ISSN 0018-9286. doi: 10.1109/TAC.1974.1100705.

[255] Stefano Palminteri, Valentin Wyart, and Etienne Koechlin. The Importance of Falsification in Computational Cognitive Modeling. *Trends in Cognitive Sciences*, 21(6):425–433, June 2017. ISSN 1364-6613. doi: 10.1016/j.tics.2017.03.011. URL http://www.sciencedirect.com/science/article/pii/S1364661317300542.

[256] Ruud Wetzels and Eric-Jan Wagenmakers. A default Bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, 19(6):1057–1064, December 2012. ISSN 1531-5320. doi: 10.3758/s13423-012-0295-x. URL https://doi.org/10.3758/s13423-012-0295-x.

[257] Benjamin Scheibehenne, Tahira Jamil, and Eric-Jan Wagenmakers. Bayesian Evidence Synthesis Can Reconcile Seemingly Inconsistent Results: The Case of Hotel Towel Reuse. *Psychological Science*, 27(7):1043–1046, 2016. ISSN 1467-9280. doi: 10.1177/0956797616644081.

[258] Robert E. Kass and Adrian E. Raftery. Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795, June 1995. ISSN 0162-1459. doi: 10.1080/01621459.1995.10476572. URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1995.10476572.

[259] Andreas Voss, Klaus Rothermund, and Jochen Brandtstädter. Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, 44(4):1048–1056, July 2008. ISSN 0022-1031. doi: 10.1016/j.jesp.2007.10.009. URL http://www.sciencedirect.com/science/article/pii/S0022103107001527.

[260] Fábio P Leite and Roger Ratcliff. What cognitive processes drive response biases? A diffusion model analysis. *Judgment and Decision Making*, 6(7):37, 2011.

[261] Martijn J. Mulder, Eric-Jan Wagenmakers, Roger Ratcliff, Wouter Boekel, and Birte U. Forstmann. Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. *Journal of Neuroscience*, 32(7):2335–2343, February 2012. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.4156-11.2012. URL http://www.jneurosci.org/content/32/7/2335.

[262] Dorion B. Liston and Leland S. Stone. Effects of Prior Information and Reward on Oculomotor and Perceptual Choices. *Journal of Neuroscience*, 28(51):13866–13875, December 2008. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3120-08.2008. URL http://www.jneurosci.org/content/28/51/13866.

[263] Kivilcim Afacan-Seref, Natalie A. Steinemann, Annabelle Blangero, and Simon P. Kelly. Dynamic Interplay of Value and Sensory Information in High-Speed Decision Making. *Current Biology*, 28 (5):795–802.e6, March 2018. ISSN 0960-9822. doi: 10.1016/j.cub.2018.01.071. URL http://www.sciencedirect.com/science/article/pii/S0960982218301040.

[264] Y. Fan, J. I. Gold, and L. Ding. Ongoing, rational calibration of reward-driven perceptual biases. *eLife*, 7, 2018. ISSN 2050-084X. doi: 10.7554/eLife.36018. URL https://europepmc.org/articles/PMC6203438/.

[265] Stephanie Goldfarb, KongFatt Wong-Lin, Michael Schwemmer, Naomi Ehrich Leonard, and Philip Holmes. Can Post-Error Dynamics Explain Sequential Reaction Time Patterns? *Frontiers in Psychology*, 3, 2012. ISSN 1664-1078. doi: 10.3389/fpsyg.2012.00213. URL https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00213/full.

[266] Ainhoa Hermoso-Mendizabal, Alexandre Hyafil, Pavel E. Rueda-Orozco, Santiago Jaramillo, David Robbe, and Jaime de la Rocha. Response outcomes gate the impact of expectations on perceptual decisions. *bioRxiv*, page 433409, May 2019. doi: 10.1101/433409. URL https://www.biorxiv.org/content/10.1101/433409v3.

[267] Jerome R. Busemeyer and James T. Townsend. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100(3):432–459, 1993. ISSN 1939-1471(Electronic),0033-295X(Print). doi: 10.1037/0033-295X.100.3.432.

[268] Edward H. Adelson and James R. Bergen. Spatiotemporal energy models for the perception of motion. *JOSA A*, 2(2):284–299, February 1985. ISSN 1520-8532. doi: 10.1364/JOSAA.2.000284. URL https://www.osapublishing.org/josaa/abstract.cfm?uri=josaa-2-2-284.

[269] Anne K. Churchland, Roozbeh Kiani, and Michael N. Shadlen. Decision-making with multiple alternatives. *Nature Neuroscience*, 11(6):693–702, June 2008. ISSN 1546-1726. doi: 10.1038/nn.2123. URL https://www.nature.com/articles/nn.2123.

[270] Paul Cisek, Geneviève Aude Puskas, and Stephany El-Murr. Decisions in changing conditions: The urgency-gating model. *The Journal of Neuroscience*, 29(37):11560–11571, 2009. ISSN 1529-2401(Electronic),0270-6474(Print). doi: 10.1523/JNEUROSCI.1844-09.2009.

[271] Xiao-Jing Wang. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968, December 2002. ISSN 0896-6273.

[272] Jiaxiang Zhang and Rafal Bogacz. Bounded Ornstein–Uhlenbeck models for two-choice time controlled tasks. *Journal of Mathematical Psychology*, 54(3):322–333, June 2010. ISSN 0022-2496. doi: 10.1016/j.jmp.2010.03.001. URL http://www.sciencedirect.com/science/article/pii/S0022249610000374.

[273] T. D. Hanks, M. E. Mazurek, R. Kiani, E. Hopp, and M. N. Shadlen. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(17):6339–6352, April 2011. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.5613-10.2011. URL https://europepmc.org/articles/PMC3356114/.

[274] Rani Moran. Optimal decision making in heterogeneous and biased environments. *Psychonomic Bulletin & Review*, 22(1):38–53, February 2015. ISSN 1531-5320. doi: 10.3758/s13423-014-0669-3. URL https://doi.org/10.3758/s13423-014-0669-3.

[275] Jan Drugowitsch, André G. Mendonça, Zachary F. Mainen, and Alexandre Pouget. Learning optimal decisions with confidence. *bioRxiv*, page 244269, April 2019. doi: 10.1101/244269. URL https://www.biorxiv.org/content/10.1101/244269v2.

[276] Amos Arieli, Alexander Sterkin, Amiram Grinvald, and Ad Aertsen. Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science*, 273(5283):1868–1871, September 1996. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.273.5283.1868. URL https://science.sciencemag.org/content/273/5283/1868.

[277] Alfonso Renart and Christian K Machens. Variability in neural activity and behavior. *Current opinion in neurobiology*, 25:211–220, April 2014.

[278] Alan E. Rorie, Juan Gao, James L. McClelland, and William T. Newsome. Integration of Sensory and Reward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP) of the Macaque Monkey. *PLoS ONE*, 5(2), February 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0009308. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2824817/.

[279] Timothy Doyeon Kim, Mohammad Kabir, and Joshua I. Gold. Coupled Decision Processes Update and Maintain Saccadic Priors in a Dynamic Environment. *Journal of Neuroscience*, 37(13):3632–3645, March 2017. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.3078-16.2017. URL http://www.jneurosci.org/content/37/13/3632.

[280] G. Pfurtscheller, A. Stancák, and C. Neuper. Post-movement beta synchronization. A correlate of an idling motor area? *Electroencephalography and Clinical Neurophysiology*, 98(4):281–293, April 1996. ISSN 0013-4694. doi: 10.1016/0013-4694(95)00258-8. URL http://www.sciencedirect.com/science/article/pii/0013469495002588.

[281] Sharareh Noorbaloochi, Dahlia Sharon, and James L. McClelland. Payoff Information Biases a Fast Guess Process in Perceptual Decision Making under Deadline Pressure: Evidence from Behavior, Evoked Potentials, and Quantitative Model Comparison. *Journal of Neuroscience*, 35(31):10989–11011, August 2015. ISSN 0270-6474, 1529-2401. doi: 10.1523/JNEUROSCI.0017-15.2015. URL http://www.jneurosci.org/content/35/31/10989.

[282] Mathieu Servant, Anna Montagnini, and Borís Burle. Conflict tasks and the diffusion framework: Insight in model constraints based on psychological laws. *Cognitive Psychology*, 72:162–195, July 2014. ISSN 0010-0285. doi: 10.1016/j.cogpsych.2014.03.002. URL http://www.sciencedirect.com/science/article/pii/S0010028514000267.

[283] Rolf Ulrich, Hannes Schröter, Hartmut Leuthold, and Teresa Birngruber. Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, 78:148–174, May 2015. ISSN 0010-0285. doi: 10.1016/j.cogpsych.2015.02.005. URL http://www.sciencedirect.com/science/article/pii/S0010028515000195.

[284] Jeffrey C Erlich, Bingni W Brunton, Chunyu A Duan, Timothy D Hanks, and Carlos D Brody. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4:e05457, April 2015. ISSN 2050-084X. doi: 10.7554/eLife.05457. URL https://elifesciences.org/articles/05457.

[285] Jochen Braun and Maurizio Mattia. Attractors and noise: Twin drivers of decisions and multistability. *NeuroImage*, 52(3):740–751, September 2010. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2009.12.126. URL http://www.sciencedirect.com/science/article/pii/S1053811910000637.

[286] Kenneth W. Latimer, Jacob L. Yates, Miriam L. R. Meister, Alexander C. Huk, and Jonathan W. Pillow. Single-trial Spike Trains in Parietal Cortex Reveal Discrete Steps During Decision-making. *Science (New York, N.Y.)*, 349(6244):184–187, July 2015. ISSN 0036-8075. doi: 10.1126/science.aaa4056. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4799998/.

[287] Hendrikje Nienborg and Bruce G. Cumming. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature*, 459(7243):89–92, May 2009. ISSN 1476-4687. doi: 10.1038/nature07821. URL https://www.nature.com/articles/nature07821.

[288] Robert Desimone and John Duncan. Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, 18(1):193–222, 1995. doi: 10.1146/annurev.ne.18.030195.001205. URL https://doi.org/10.1146/annurev.ne.18.030195.001205.

[289] John H. Reynolds and David J. Heeger. The Normalization Model of Attention. *Neuron*, 61(2):168–185, January 2009. ISSN 0896-6273. doi: 10.1016/j.neuron.2009.01.002. URL http://www.sciencedirect.com/science/article/pii/S0896627309000038.

[290] Valentin Wyart, Anna Christina Nobre, and Christopher Summerfield. Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences*, 109(9):3593–3598, February 2012. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1120118109. URL https://www.pnas.org/content/109/9/3593.

[291] Braden A. Purcell and Roozbeh Kiani. Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proceedings of the National Academy of Sciences*, 113(31):E4531–E4540, August 2016. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1524685113. URL https://www.pnas.org/content/113/31/E4531.

[292] Christopher J Honey, Thomas Thesen, Tobias H Donner, Lauren J Silbert, Chad E Carlson, Orrin Devinsky, Werner K Doyle, Nava Rubin, David J Heeger, and Uri Hasson. Slow cortical dynamics and the accumulation of information over long timescales. *Neuron*, 76(2):423–434, October 2012.

[293] John D. Murray, Alberto Bernacchia, David J. Freedman, Ranulfo Romo, Jonathan D. Wallis, Xinying Cai, Camillo Padoa-Schioppa, Tatiana Pasternak, Hyojung Seo, Daeyeol Lee, and Xiao-Jing Wang. A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience*, 17(12):1661–1663, December 2014. ISSN 1546-1726. doi: 10.1038/nn.3862. URL https://www.nature.com/articles/nn.3862.

[294] Rishidev Chaudhuri, Kenneth Knoblauch, Marie-Alice Gariel, Henry Kennedy, and Xiao-Jing Wang. A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*, 88(2):419–431, October 2015. ISSN 0896-6273. doi: 10.1016/j.neuron.2015.09.008. URL http://www.sciencedirect.com/science/article/pii/S0896627315007655.

[295] Caroline A. Runyan, Eugenio Piasini, Stefano Panzeri, and Christopher D. Harvey. Distinct timescales of population coding across cortex. *Nature*, 548(7665):92–96, August 2017. ISSN 0028-0836. doi: 10.1038/nature23020. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5859334/.

[296] Mark E. Mazurek, Jamie D. Roitman, Jochen Ditterich, and Michael N. Shadlen. A Role for Neural Integrators in Perceptual Decision Making. *Cerebral Cortex*, 13(11):1257–1269, November 2003. ISSN 1047-3211. doi: 10.1093/cercor/bhg097. URL https://academic.oup.com/cercor/article/13/11/1257/274091.

[297] Harold Stanislaw and Natasha Todorov. Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1):137–149, March 1999. ISSN 1532-5970. doi: 10.3758/BF03207704. URL https://doi.org/10.3758/BF03207704.

[298] Roger Ratcliff and Francis Tuerlinckx. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9(3):438–481, September 2002. ISSN 1531-5320. doi: 10.3758/BF03196302. URL https://doi.org/10.3758/BF03196302.

[299] Andrei R. Teodorescu, Rani Moran, and Marius Usher. Absolutely relative or relatively absolute: Violations of value invariance in human decision making. *Psychonomic Bulletin & Review*, 23(1):22–38, 2016. ISSN 1531-5320(Electronic),1069-9384(Print). doi: 10.3758/s13423-015-0858-8.

[300] James H. Steiger. Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2):245–251, 1980. ISSN 1939-1455(Electronic),0033-2909(Print). doi: 10.1037/0033-2909.87.2.245.

[301] S L Rogers and L T Friedhoff. Pharmacokinetic and pharmacodynamic profile of donepezil HCl following single oral doses. *British Journal of Clinical Pharmacology*, 46(Suppl 1):1–6, November 1998. ISSN 0306-5251. doi: 10.1046/j.1365-2125.1998.0460s1001.x. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1873812/.

[302] John-Michael Sauer, Barbara J. Ring, and Jennifer W. Witcher. Clinical Pharmacokinetics of Atomoxetine. *Clinical Pharmacokinetics*, 44(6):571–590, June 2005. ISSN 1179-1926. doi: 10.2165/00003088-200544060-00002. URL https://doi.org/10.2165/00003088-200544060-00002.

[303] Samuel R. Chamberlain, Adam Hampshire, Ulrich Müller, Katya Rubia, Natalia del Campo, Kevin Craig, Ralf Regenthal, John Suckling, Jonathan P. Roiser, Jon E. Grant, Edward T. Bullmore, Trevor W. Robbins, and Barbara J. Sahakian. Atomoxetine Modulates Right Inferior Frontal Activation During Inhibitory Control: A Pharmacological Functional Magnetic Resonance Imaging Study. *Biological Psychiatry*, 65(7):550–555, April 2009. ISSN 0006-3223. doi: 10.

1016/j.biopsych.2008.10.014. URL http://www.sciencedirect.com/science/article/pii/S0006322308012432.

[304] Ariel Rokem and Michael A. Silver. Cholinergic Enhancement Augments Magnitude and Specificity of Visual Perceptual Learning in Healthy Humans. *Current Biology*, 20(19):1723–1728, October 2010. ISSN 0960-9822. doi: 10.1016/j.cub.2010.08.027. URL http://www.sciencedirect.com/science/article/pii/S0960982210010171.