CHOICE HISTORY BIASES SUBSEQUENT EVIDENCE ACCUMULATION

Anne E. Urai[1,2,3,*], Jan Willem de Gee[1,2], Konstantinos Tsetsos[1,#], Tobias H. Donner[1,2,4,#,*]

1 Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

2 Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

3 Current address: Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA

4 Amsterdam Brain & Cognition, University of Amsterdam, Amsterdam, The Netherlands

Shared senior authorship

* Correspondence: anne.urai@gmail.com, t.donner@uke.de

ABSTRACT

Perceptual choices depend not only on the current sensory input, but also on the behavioral context. An important contextual factor is the history of one's own choices. Choice history often strongly biases perceptual decisions, and leaves traces in the activity of brain regions involved in decision processing. Yet, it remains unknown how such history signals shape the dynamics of later decision formation. Models of perceptual choice construe decision formation as the accumulation of sensory evidence towards decision bounds. In this framework, it is commonly assumed that choice history signals shift 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 behavioral data from perceptual choice tasks, 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.

INTRODUCTION

Decisions are not isolated events, but are embedded in a sequence of choices. Preceding choices can exert a large influence even on low-level perceptual judgments (Fernberger, 1920; Treisman and Williams, 1984). Previous work on the mechanisms of perceptual decision-making has largely focused on processing of the current sensory evidence (Gold and Shadlen, 2007). Yet, humans (Fründ et al., 2014; Urai et al., 2017), monkeys (Gold et al., 2008) and rodents (Busse et al., 2011; Odoemene et al., 2018) base their choices not only on current sensory input, but also on choice history, in a way that can be adjusted to serial correlations in the sensory environment (Abrahamyan et al., 2016; Kim et al., 2017; Braun et al., 2018). History biases are also prevalent in environments lacking such structure, and vary substantially across individuals (Fründ et al., 2014; Abrahamyan et al., 2016; Urai et al., 2017). Choice history biases are ubiquitous also in other domains of decision-making (Leopold et al., 2002; Allefeld et al., 2013; Padoa-Schioppa, 2013), likely reflecting a general principle of the mechanisms governing decision-making.

Computational theory (Gao et al., 2009; Glaze et al., 2015) and psychophysical data (Kim et al., 2017; Braun et al., 2018) indicate that choice history biases result from the accumulation of internal decision variables across trials, with a timescale governed by the decision-makers' internal model of the correlation structure of their environment. Neural signals reflecting previous choices have been found across the sensorimotor pathways of the cerebral cortex, from sensory to associative and motor regions (de Lange et al., 2013; Akaishi et al., 2014; Pape and Siegel, 2016; Purcell and Kiani, 2016a; St. John-Saaltink et al., 2016; Thura et al., 2016; Hwang et al., 2017; Scott et al., 2017).

By which mechanism are choice history signals incorporated into the formation of a decision? Current models of perceptual decision-making posit the accumulation of noise-corrupted sensory evidence over time, resulting in an internal decision variable that grows with time (Bogacz et al., 2006; Gold and Shadlen, 2007; Ratcliff and McKoon, 2008; Brody and Hanks, 2016). 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 be brought about 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 accumulates (Figure 1). The former mechanism can be thought of as adding an offset to the perceptual interpretation of the current sensory evidence during the generation of the response. The latter mechanism corresponds to biasing that perceptual interpretation itself, analogous to selective attention.

Existing computational models of choice history biases postulate a shift in the starting point of the decision variables towards the bound of the previous choice (Yu and Cohen, 2008; Zhang et al., 2014; Glaze et al., 2015). This is a natural prediction, assuming that history biases are due to the slow (passive) decay of the decision variable into the next trial (Cho et al., 2002; Gao et al., 2009, 2009; Bonaiuto et al., 2016). However, the cerebral cortex is equipped with a hierarchy of timescales (Honey et al., 2012; Murray et al., 2014; Chaudhuri et al., 2015; Runyan et al., 2017; Scott et al., 2017). Choice history biases might originate from the slow (i.e., tens of seconds) across-trial accumulation of decision variables, at higher stages of the hierarchy than those accumulating current evidence on timescales of hundreds of milliseconds. Previous experimental work quantifying history biases in perceptual choice either did not analyze the within-trial dynamics of decision formation (Busse et al., 2011; de Lange et al., 2013; Akaishi et al., 2014; Fründ et al., 2014; Urai et al., 2017; Braun et al., 2018), or did not allow for distinguishing between starting point and accumulation biases (Cho et al., 2002; Gold et al., 2008; Yu and Cohen, 2008; Gao et al., 2009; Wilder et al., 2009; Bode et al., 2012; Jones et al., 2013; Zhang et al., 2014). Consequently, it remains unknown which of those two computational mechanisms accounts for the choice history biases observed in overt behavior.

Here, we addressed this issue by fitting a bounded-accumulation decision models to human behavioral data from perceptual choice tasks. Across a variety of task protocols and sensory modalities, we found that history biases in individual choice behavior were consistently explained by a history-dependent variation in the accumulation bias, rather than the starting point. This indicates that the interaction between choice history and decision formation is more complex than previously thought: internal decision signals may act like a cue for selective attention, which then biases evidence accumulation towards (or away from) the previous chosen perceptual interpretation of the sensory input.

RESULTS

We fit the drift diffusion model to behavioral data (choices and response times, RT) from a total of 194 human participants (Materials and Methods). The drift diffusion model (DDM) is one variant of bounded-accumulation models, popular because it provides good fits to RT and choice patterns from a large array of two-choice task (Ratcliff and McKoon, 2008) and its latent variables (model parameters) seem to correspond to neural signals observed at different stages of the sensory-motor pathways of the brain (Gold and Shadlen, 2007). The shape of RT distributions and choice fractions jointly constrain the model parameters. We here 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 (Supplementary Figure 1). We refer to the latter parameter (termed "drift criterion" by Ratcliff and McKoon, 2008) as "drift bias".

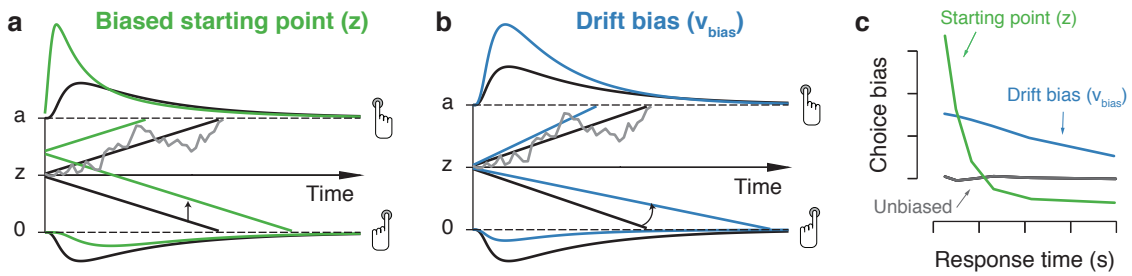


Figure 1. Two biasing mechanisms within the drift diffusion model. *The model postulates that noisy sensory evidence is accumulated over time, until the resulting decision variable y reaches one of two bounds (dashed black lines at y=0 and y=a) for the two choice options. Repeating this process over 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. (**b**) Choice history-dependent shift in drift bias. Blue lines: mean drift and RT distributions under biased drift. Both mechanisms differentially affect the shape of RT distributions. (**c**) Conditional bias functions (White and Poldrack, 2014), showing the fraction of biased choices as a function of RT.*

Within the DDM, choice behavior can be biased by two mechanisms: shifting the starting point or biasing the drift towards one of the bounds (Figure 1). A shift in starting point would be 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). These biasing mechanisms are hard to differentiate based on the proportion of choices alone, but they are readily distinguishable by the shape of RT distributions or the relationship between choice bias and RT (Figure 1c).

We fit different variants of the DDM to data from six different experiments. These covered a range of task protocols and sensory modalities commonly used in studies of perceptual decision-making (see 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 (Abrahamyan et al., 2016; Urai et al., 2017), 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. Several observations indicate that the DDM provided fit the data well overall (Supplementary Figure 2). First, RT distributions matched the model predictions reasonably well (shown separately for each combination of stimuli and choices in Supplementary Figure 2b, darker colors indicate predicted RTs obtained through model simulations). Second, for the fits obtained with a hierarchical Bayesian fitting procedure (see Supplementary Figure 1 and Materials and Methods), used for Figures 3-5, the $\widehat{R}$ 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', Supplementary Figure 2a) and monotonically increased with stronger sensory evidence (Supplementary Figure 2a). In fixed duration tasks, the decision-maker does not need to set a bound for terminating the decision (Bogacz et al., 2006), 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 (e.g. Ratcliff, 2006; Bode et al., 2012; Jahfari et al., 2012) and might have reflected the fact that observers set implicit decision bounds also when they do not control the stimulus duration (Kiani et al., 2008; but see Tsetsos et al., 2015).

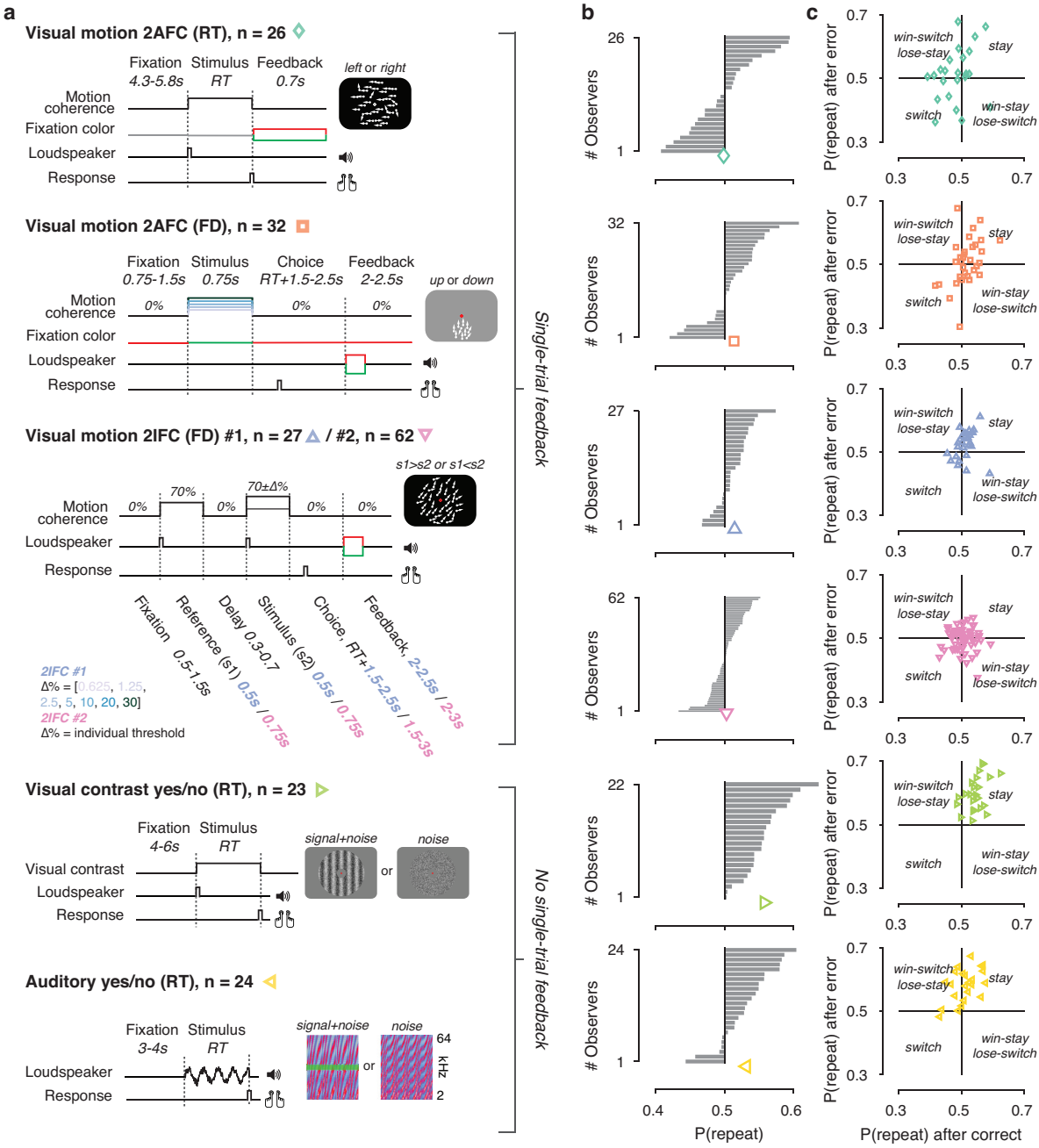


Figure 2. Behavioral tasks and individual differences. *(**a**) Schematics of perceptual decision-making tasks used in each dataset. See Methods section* Datasets: behavioral tasks and participants. *(**b**) Distributions of individual choice history biases for each dataset. Grey bars show individual 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.*

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

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 (Supplementary Figure 3). 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 considerable dependence on previous choices (Figure 2f). The model with both history-dependent starting point and drift bias provided the best fit to all six datasets (Figure 3a), based on the Deviance Information Criterion (DIC).

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 those diagnostic features in the data (Palminteri et al., 2017) that distinguished starting point from drift bias. As shown above (Figure 1), 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 four different model variants and computed so-called 'conditional bias functions' (White and Poldrack, 2014): 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 bars).

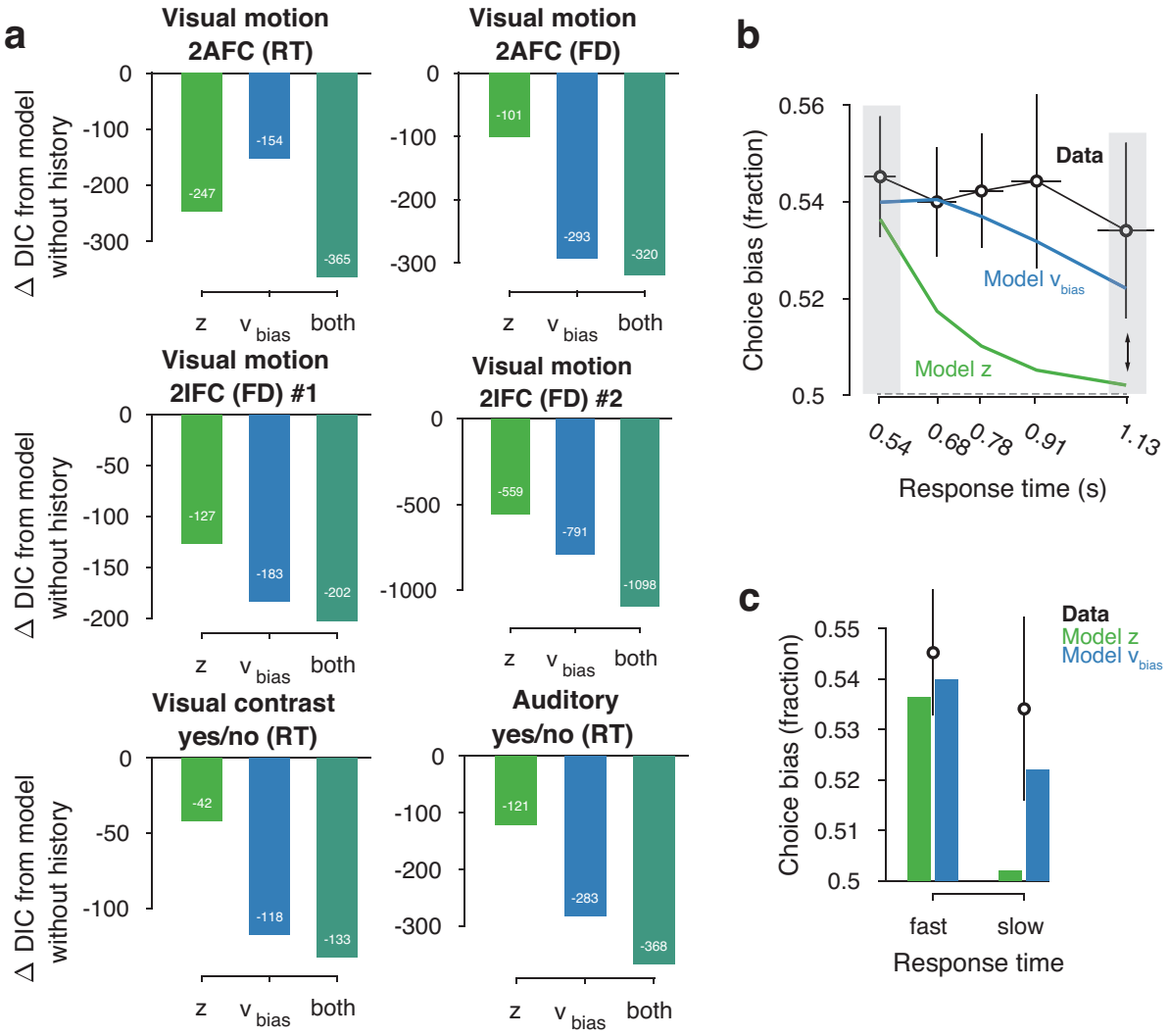


Figure 3. Model comparison and simulations. *(**a**) For each dataset, we compared the DIC between models where drift bias, starting point bias or both were allowed to vary as a function of previous choice. The DIC for a model without history dependence was used as a baseline for each dataset. Lower DIC values indicate a model that is better able to explain the data, after taking into account the model complexity; a ΔDIC 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 and drift bias 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 ± 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.*

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. 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, Δr values).

We quantified the total evidence by computing a Bayes factor for each correlation (Wetzels and Wagenmakers, 2012), and multiplying these across datasets (Scheibehenne et al., 2016). 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 (Kass and Raftery, 1995).

The same qualitative pattern of results was obtained with an alternative fitting procedure (non-hierarchical G^2 optimization, Supplementary Figure 4a), as well as a model that allowed for additional across-trial variability in non-decision time (Supplementary Figure 4b). 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 (Supplementary Figure 5).

The lack of a correlation between history-dependent starting point shifts and individual choice repetition is surprising in light of previous accounts (Yu and Cohen, 2008; Gao et al., 2009; Glaze et al., 2015). 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 (Supplementary Figure 3, 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. Critically, only the history-dependent shift in drift accounted for individual differences in choice repetition (Figure 4).

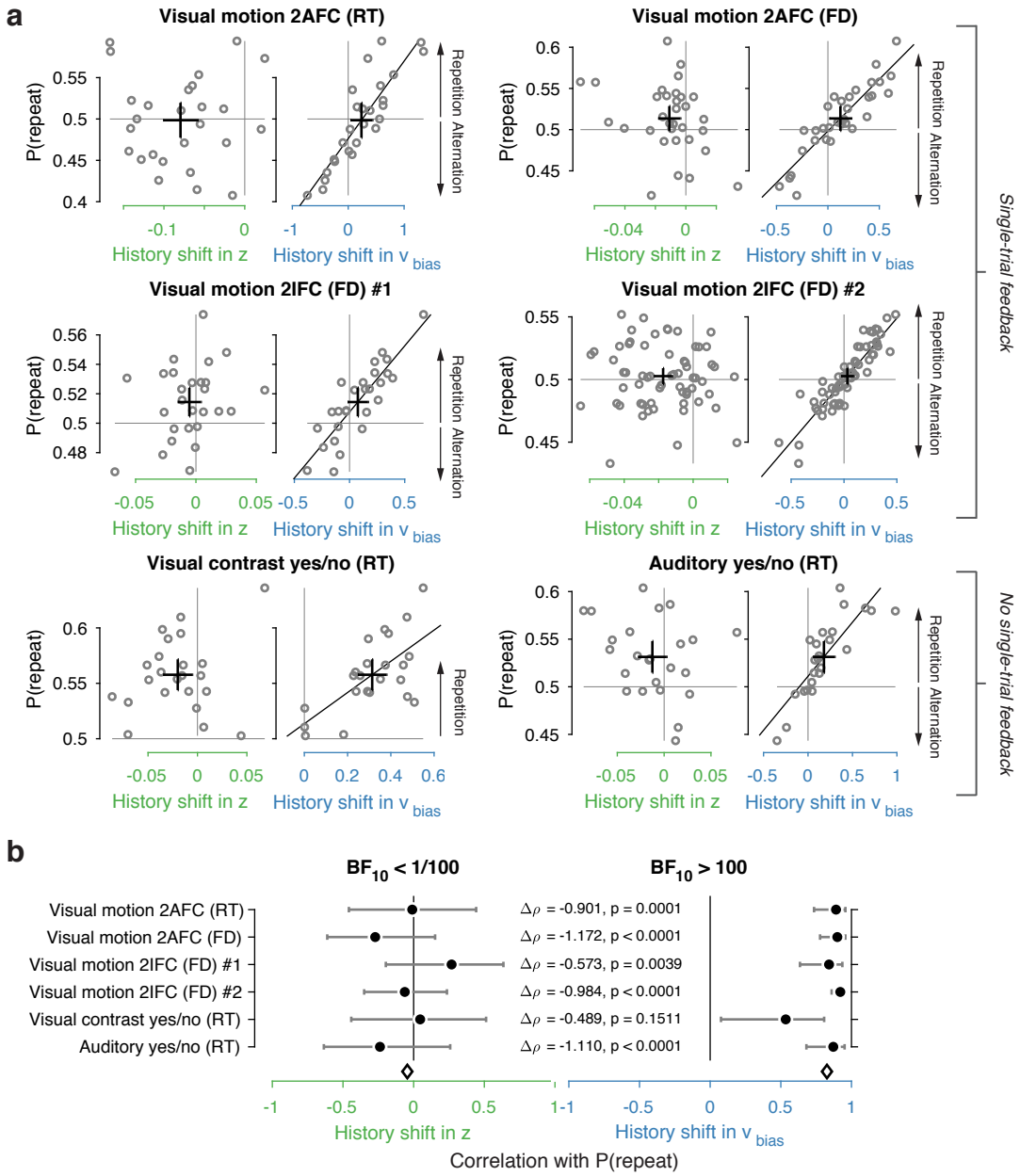


Figure 4. Individual choice history biases are explained by history-dependent changes in drift bias, not starting point. *(**a**) Relationship between individual choice repetition probabilities, P(repeat), and history shift in starting point (left column, green) and drift (right column, blue). Parameter 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 ρ) 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. Δρ 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.

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 (Voss et al., 2008; Leite and Ratcliff, 2011; Mulder et al., 2012), but might also bias drift (Liston and Stone, 2008; Afacan-Seref et al., 2018). 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 results hold 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, when separately estimating all model parameters (history shift in starting point and drift bias) and model-free measures (choice repetition probability) after correct and error trials, we found similar results after both (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').

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 non-specific sequential effects (Supplementary Figure 6) in line with previous work on post-error slowing (Dutilh et al., 2012; Goldfarb et al., 2012; Purcell and Kiani, 2016a).

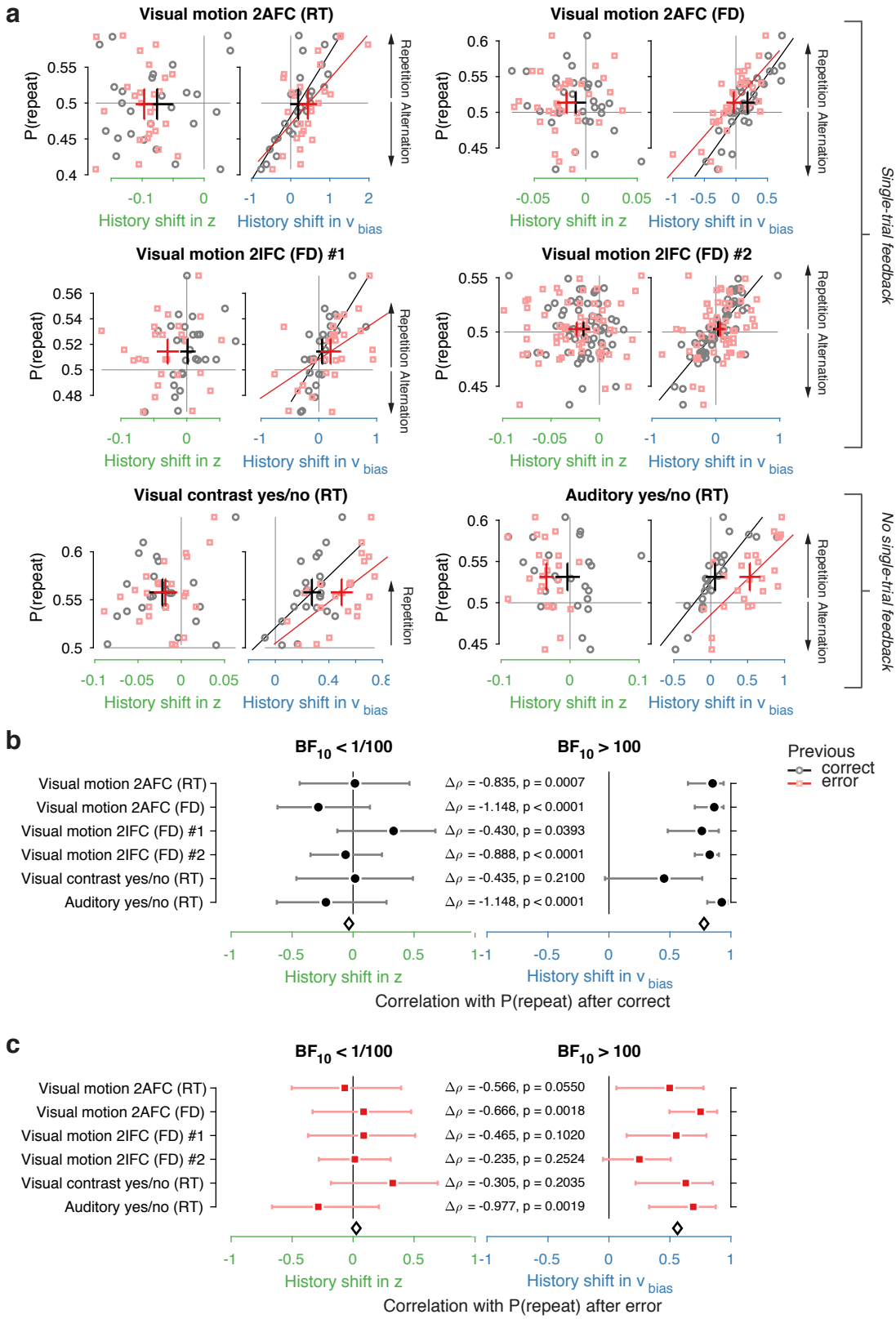


Figure 5. History shift in drift bias explains individual choice behavior after both error and correct decisions. *As in Figure 4, but separately following correct (black) and error (red) trials. (**a**) Relationship between repetition probability and history shifts in starting point and drift bias, separately computed for trials following correct (black circles) and error (red squares) responses. (**b**)*

*Summary of correlations (as in Figure 4c) for trials following a correct response. (**c**) Summary of correlations (as in Figure 4c) for trials following an error response.*

HISTORY-DEPENDENT ACCUMULATION BIAS EXPLAINS INDIVIDUAL CHOICE REPETITION BEHAVIOR IRRESPECTIVE OF SPECIFICS OF BOUNDED-ACCUMULATION MODELS

To establish the generality of our conclusions and further pinpoint 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. These models included variants of the DDM (i.e. perfect accumulator) with more complex dynamics of the bias or the decision bounds, as well as variants of a leaky accumulator (Busemeyer and Townsend, 1993; Usher and McClelland, 2001; Brunton et al., 2013). 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 (Adelson and Bergen, 1985; Urai and Wimmer, 2016; 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 (Brunton et al., 2013).

We first re-fit the standard DDM with the two biasing parameters allowed to vary with previous choice (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 6b1). 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 BIC). 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 (Afacan-Seref et al., 2018). 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 6a, 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 6b2). Furthermore, in the latter (hybrid) model, the

constant drift bias more strongly predicted individual choice repetition behavior (Figure 6c2), in line with the constant accumulation bias inferred from the standard DDM fits. For the fits shown in Figure 6 b2/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 time-average 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, as used for the more complex dynamical models described below (ΔBIC relative to no-history model: -766, -658, -341, for constant drift bias, ramping drift bias, and hybrid, respectively; correlation with P(repeat): $\rho(30) = 0.5458$, $p = 0.0012$; $\rho(30) = 0.3600$, $p = 0.0429$ for constant and ramping drift bias, respectively).

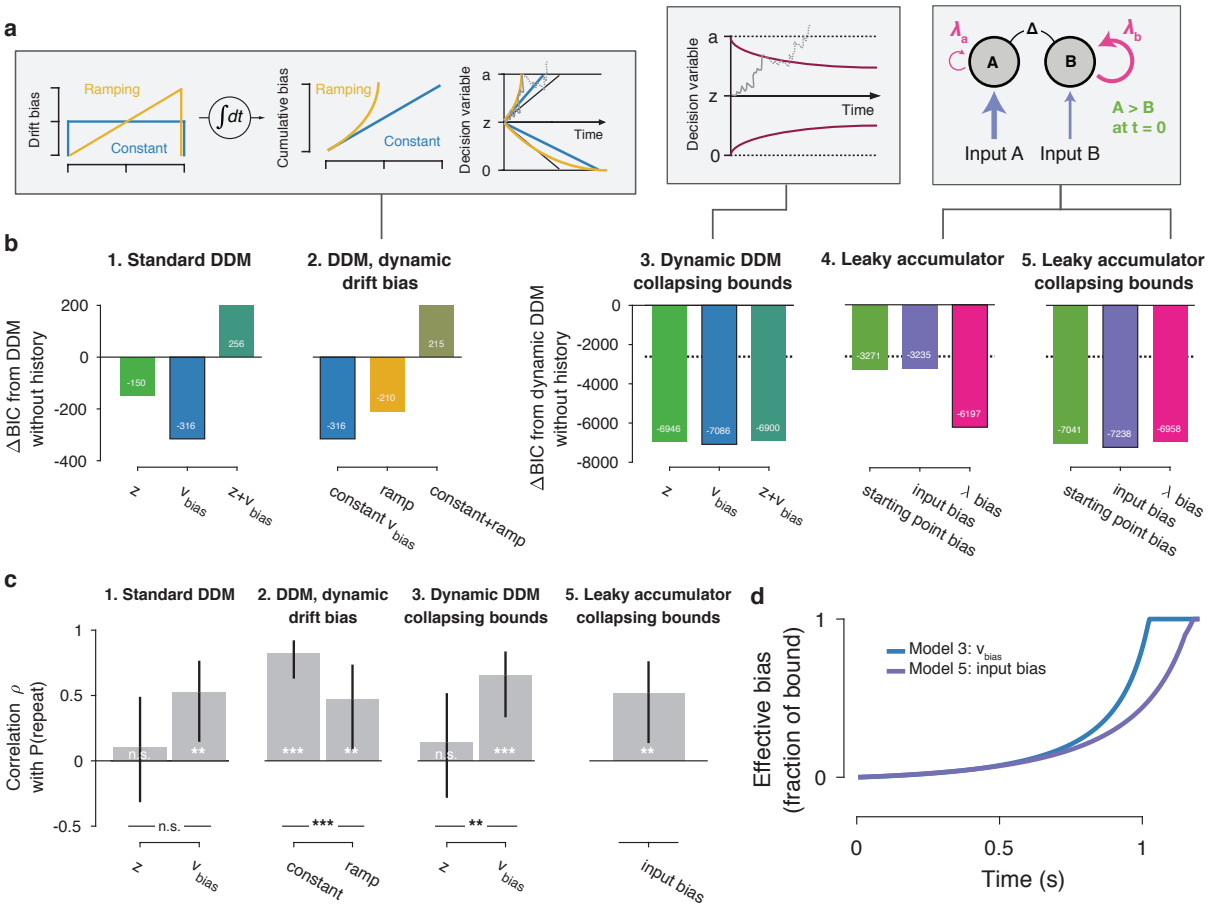


Figure 6. Extended dynamic models of biased evidence accumulation. *(**a**) Model schematics. See main text for details. In the third panel from the left, the stimulus-dependent mean drift is shown in black, overlaid by the biased mean drift in color (as in Figure 1a,b). **(b)** BIC values for each history-dependent model, as compared to a standard (left) or dynamic (right) DDM without history. The winning model (lowest BIC value) within each model class is shown with a black outline. In 3-5, the*

dashed line indicates the BIC 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.*

It has been proposed that decision bounds might collapse over time, implementing an 'urgency signal' (Figure 6a, middle; Churchland et al., 2008; Cisek et al., 2009). Indeed, adding collapsing bounds substantially improved our model fits (Figure 6b3). 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 6b3) and captured individual choice repetition behavior (Figure 6c3) 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 (Usher and McClelland, 2001; Wang, 2002). Under certain parameter regimes, these circuit dynamics can be reduced to lower-dimensional models (Bogacz et al., 2006; Wong and Wang, 2006). 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 (Usher and McClelland, 2001). Evidence accumulation can then be biased through an internal representation of the sensory input or through the way this sensory representation is accumulated (Figure 6a, right). We here used a reduced competing accumulator model, where the decision variable was computed as the difference of two leaky accumulators (Busemeyer and Townsend, 1993; Zhang and Bogacz, 2010; see also Brunton et al., 2013) 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 6a, 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 λ, without implying that $\lambda<0$ (corresponding to activation decay). Our fits allowed λ to take either negative ('forgetful' regime) or positive ('unstable' regime) values (Supplementary Figure 7d; see also Brunton et al, 2013). 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 (Supplementary Figure 8g, h).

Fitting different variants of these leaky accumulator models to participants' data showed that adding a collapsing bound substantially improved the fits, just as in the DDM (Figure 6b4, 5, see also Supplementary Figure 7d). Critically, the leaky accumulator with biased accumulator *input* best explained the data, among all the models considered (Figure 6b4). Furthermore, the individually estimated input bias predicted individual choice repetition (Figure 6c5). This suggests that choice history might specifically bias the internal representation of the sensory evidence feeding into the evidence accumulation process.

DYNAMICS OF EFFECTIVE BIAS SIGNAL APPROXIMATES RATIONAL COMBINATION OF PRIOR INFORMATION WITH 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 (Hanks et al., 2011). We simulated this signal based on the group average parameters for the best-fitting DDM and leaky accumulator models (Figure 6d). In the DDM and leaky accumulator (both with collapsing bound), the effective bias accelerated.

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 6a, left). Thus, the observation that the latter model lost by a wide margin against the two models with more complex

dynamics (Figure 6b, 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 6b5), 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.

Taken together, 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 the previous 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 (Hanks et al., 2011; Moran, 2015). 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 (Hanks et al., 2011). Indeed, within the framework of the DDM, this way of combining prior information with current evidence maximizes reward rate (Moran, 2015; see also Drugowitsch and Pouget, 2018). 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 (Moran, 2015), due to variations in the external input (i.e., mean drift rate in the DDM) or internal factors (i.e., drift rate variability in the DDM), such as the inherent variability of sensory cortical responses (Arieli et al., 1996; Faisal et al., 2008). In particular, the dataset from Figure 6 entailed strong trial-to-trial variations in the external input (Supplementary Figure 7a,b). Thus, the dynamics of the effective bias signal uncovered in Figure 6d suggest that participants combined prior information with current evidence in a rational fashion.

DISCUSSION

Observers' perceptual choices often depend on choices made before: how is such history information incorporated into the formation of the next decision? Fitting bounded accumulation decision models to behavioral data from a range of psychophysical experiments, we teased apart two possible effects of choice history on the current decision process: a shift in the starting point of the decision variable, or a change in the rate at which evidence for one versus the other option is accumulated. These two scenarios can lead to identical (history-dependent) asymmetries in choice fractions, but can be distinguished based on their effects on RT distributions. We obtained consistent evidence for the second scenario: history shifts in accumulation bias are the dominant source of individual variability in choice history biases. This result calls for a revision of current models of choice history biases (Yu and Cohen, 2008; Zhang et al., 2014).

Previous quantitative treatments of perceptual decision-making have commonly ignored endogenously generated biases. Choice history biases vary dynamically from trial to trial as a function of the history of previous choices (and possibly other experimental events, such as choice outcomes). Ignoring choice history leads to underestimation of the observer's true perceptual sensitivity (Fründ et al., 2014; Abrahamyan et al., 2016; Braun et al., 2018). More generally, trial-to-trial behavioral variability resulting from dynamic internal biases has commonly been attributed to 'noise' in the decision computation (Shadlen et al., 1996; Renart and Machens, 2014; Wyart and Koechlin, 2016). Our results imply that choice history accounts for a significant portion of the trial-to-trial variability in evidence accumulation. Tracking internal factors in perceptual decision-making, such as choice history, allows for partitioning behavioral variability into systematic components, and residual variability that might reflect noise at the level of the underlying neural computations.

Our results are in line with the observation that for the DDM in heterogeneous environments (i.e., variations in evidence reliability), reward is maximized by incorporation prior information into both starting point and drift bias (Moran, 2015). It is instructive to relate our results with previous studies that manipulated choice bias through asymmetries in prior probability or reward magnitude. Such biases have mostly been associated with shifts in starting point (Leite and Ratcliff, 2011; Mulder et al., 2012; White and Poldrack, 2014; Rorie et al., 2010; Gao et al., 2011; but only for decisions without time pressure, see Afacan-Seref et al., 2018). Yet, one study provided clear evidence for an accumulation bias induced by manipulations in prior probability, under variations of external evidence strength (Hanks et al., 2011). Accumulation biases could also be induced through variations in

arbitrary boundaries between stimulus categories - for example, whether there are 'many' or 'few' items in a display (Leite and Ratcliff, 2011; White and Poldrack, 2014). The biases studied in this previous work were under experimental control, whereas the choice history biases we studied emerge spontaneously and in an idiosyncratic fashion (Figure 2e).

The question of *how* prior information is combined with new evidence during decision formation is distinct from the question of *why* participants used choice history as a prior for their decisions. In our experiments the sensory evidence was by design uncorrelated across trials (as in the vast majority of studies into perceptual decision-making). Thus, any history bias can only reduce performance below the level that could be achieved given the observer's evidence sensitivity. One possible explanation is that decision-makers learn the stability of real-world sensory evidence (auto-correlated across various timescales) and erroneously apply this model to randomized laboratory experiments (Yu and Cohen, 2008). This internal model of the environmental stability is the so-called subjective hazard rate, the inferred probability of changes in the sources generating sensory evidence (Yu and Cohen, 2008; Glaze et al., 2015). Experimental evidence shows that hazard rate governs the adaptive use of choice history information (as a prior that evolves slowly across trials) in human perceptual decision-making (Glaze et al., 2015; Kim et al., 2017; Braun et al., 2018). Consequently, any bias in the subjective hazard rate (i.e., a systematic deviation from the objective hazard rate, which equals 0.5 in our experiments) will produce maladaptive choice history biases: it will push decision-makers towards choice repetition or alternation. Subjective hazard rates might be biased, for example, through an erroneous inference about the environmental stability. Under natural conditions, not only the sources of the evidence may change, but also the rate of changes between these sources (i.e. hazard rate). Along with their history prior about the current source, agents may also continuously update their subjective hazard rates based on sequences of choices, evidence, and/or feedback (Yu and Dayan, 2005; Mathys et al., 2014; Meyniel et al., 2016; Glaze et al., 2018; Hermoso-Mendizabal et al., 2018). In this framework, the individual differences in choice history biases that we exploited in this study (i.e., correlation analyses) can be accounted for by individual differences in subjective hazard rates. Combined with our conclusions from the time course of the effective bias signal (see Results section *Dynamics of the effective bias signal approximate rational combination of prior information with current evidence*), these considerations suggest that participants may have applied a rational strategy, but based on erroneous assumptions about the structure of the environment.

The conclusion that choice history biases evidence accumulation provides a better explanation of our results than several possible 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 (Noorbaloochi et al., 2015), or engage in automatic decision processing (Servant et al., 2014; Ulrich et al., 2015) or post-accumulation biases (Erlich et al., 2015). Also, the decision computation may entail noise-driven attractor dynamics (Wang, 2002; Braun and Mattia, 2010) possibly with sudden 'jumps' between neural activity states (Latimer et al., 2015), or probability summation across time (Watson, 1979) 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 (see Supplementary Figure 8 for simulations). For example, in the standard DDM the combination of 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 (Supplementary Figure 8, compare b and c). We are not aware of any empirical evidence for such asymmetric changes in decision bounds. Preparatory neural activity in association and motor cortex seems to always reach a fixed threshold level just before response, irrespective of the prior probability of choice options (Hanks et al., 2011) or the speed-accuracy trade-off (Hanks et al., 2014; Murphy et al., 2016), while the build-up of this activity is biased by prior information (Hanks et al., 2011). On the other hand, a large body of work on selective attention has established that the brain is equipped with powerful machinery by which neural representations of sensory evidence are biased in a top-down fashion (Reynolds and Heeger, 2009). Attentional cues induce top-down signals in higher-level cortical regions that are also implicated in evidence accumulation. Those top-down signals then feed back across the cortical hierarchy to bias representations of sensory evidence in visual cortex towards specific features of the input (Desimone and Duncan, 1995). Choice history signals could act as sources of this top-down modulation of sensory cortex (Nienborg and Cumming, 2009; Wimmer et al., 2015; St. John-Saaltink et al., 2016; but see Gold et al., 2008) In sum, our findings are in line with the idea that choice

history signals may act as internal 'cues' that direct (feature-based) attention towards or away from previously chosen options.

A growing body of evidence suggests that perceptual decision-making arises from accumulation of neural signals over multiple timescales. Behavioral analyses have revealed effective (within-trial) evidence accumulation over timescales ranging from a few hundred milliseconds (Kiani et al., 2008; Tsetsos et al., 2015) to several seconds (Tsetsos et al., 2012; Wyart et al., 2012; Cheadle et al., 2014). Further, humans can flexibly accumulate the same sensory input with different timescales, dependent on their model of environmental statistics (Ossmy et al., 2013; Glaze et al., 2015). They can accumulate internal signals (decision variables) across trials, over tens of seconds to build up choice history biases (Abrahamyan et al., 2016; Purcell and Kiani, 2016b; Braun et al., 2018). This behavioral evidence sits well with hierarchy of intrinsic timescales of cortical circuits inferred from neural dynamics in different cortical regions (Honey et al., 2012; Murray et al., 2014; Chaudhuri et al., 2015; Runyan et al., 2017; Scott et al., 2017). It is tempting to speculate that the history-dependent evidence accumulation biases we have uncovered here result from long timescale (across-trial) accumulators at higher stages, that are also involved in the control of attention. These high-level accumulators, then bias low-level sensory representations, in turn giving rise to evidence accumulation biases at intermediate stages of the cortical hierarchy.

MATERIALS AND METHODS

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 Mazurek et al. (2003). These two protocols have also been termed 'free response protocol' and 'interrogation protocol' (Bogacz et al., 2006). 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 DIRECTION DISCRIMINATION TASK (RT)

These data were previously published (Murphy et al., 2014), and are available at https://doi.org/10.5061/dryad.tb542. 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 DIRECTION DISCRIMINATION TASK (FD)

Participants and informed consent. 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.

Task and procedure. 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/degree^2, 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*; Supplementary Figure 7a,b). 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)'.

VISUAL MOTION COHERENCE DISCRIMINATION 2IFC TASK (FD): DATASET 1

These data were previously published in (Urai et al., 2017), and are available at http://dx.doi.org/10.6084/m9.figshare.4300043. 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'.

2IFC VISUAL MOTION COHERENCE DISCRIMINATION TASK (FD): DATASET 2

Participants and informed consent. Sixty-two participants (aged 19-35 years, 43 women and 19 men) participated in the study after screening for psychiatric, neurological 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.

Task protocol. 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 (Supplementary Figure 5b). We refer to this task as 'Visual motion 2IFC (FD) #2'.

VISUAL CONTRAST YES/NO DETECTION TASK (RT)

These data were previously published (de Gee et al., 2014), and are available at https://doi.org/10.6084/m9.figshare.4806559. 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 (de Gee et al., 2014) 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 (de Gee et al., 2017), and are available at https://doi.org/10.6084/m9.figshare.4806562. 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; McGinley et al., 2015) 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)'.

MODEL-FREE ANALYSIS OF SENSITIVITY AND CHOICE HISTORY BIAS

We quantified perceptual sensitivity in terms of signal detection-theoretic d' (Green and Swets, 1966):

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

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 (Stanislaw and Todorov, 1999). 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.

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 (Supplementary Figure 2b).

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 (Palminteri et al., 2017). Specifically, we computed conditional bias functions (White and Poldrack, 2014) 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). To generate 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. This allowed us to visualize the effect of choice history bias as a function of time within each trial, together for all observers in each dataset. The shape of the conditional bias functions for models with z and v_{bias} confirm that z predominantly produces biased choices with short RTs, whereas v_{bias} leads to biased choices across the entire range of RTs (Figure 3b).

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 (Ratcliff and McKoon, 2008). Ignoring non-decision time, drift rate variability, and starting point variability (see below), the DDM describes the accumulation of noisy sensory evidence as follows:

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

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 (Bogacz et al., 2006). In an unbiased case, the starting point of the decision variably $\mathrm{y}(0) = \mathrm{z}$, is situated midway between the two decision bounds 0 and a:

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

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

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

A drift bias can be implemented by adding a stimulus-independent constant $\mathrm{v}_{\mathrm{bias}}$, also referred to as drift bias (Ratcliff and McKoon, 2008), to the (stimulus-dependent) mean drift (Figure 1b). This adds a bias to the drift that linearly grows with time:

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

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, $\mathrm{z}_{\mathrm{bias}}$ and $\mathrm{v}_{\mathrm{bias}}$ have also been referred to as 'prior' and 'dynamic' bias (Moran, 2015) or 'judgmental' and 'perceptual' bias (Liston and Stone, 2008).

ESTIMATING HDDM BIAS PARAMETERS

We used hierarchical drift diffusion modeling as implemented in the HDDM toolbox (Wiecki et al., 2013) 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 (Ratcliff and Tuerlinckx, 2002). 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 (Supplementary Figure 1). Across-trial variability in drift rate and starting point were estimated at the group-level only (Ratcliff and Childers, 2015). 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; cf Gao et al., 2009; Wilder et al., 2009). 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 MODEL FITTING PROCEDURES

The HDDM (Wiecki et al., 2013) 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 (see e.g. Supplementary Supplementary Figure 3). 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 (Supplementary Supplementary Figure 1; Wiecki et al., 2013). 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 G^2 optimization (Ratcliff and Tuerlinckx, 2002) and obtained qualitatively identical results (Supplementary Figure 4a).

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 $\widehat{R}$ statistic (which compares within-chain and between-chain variance) and checked that all group-level parameters had an $\widehat{R}$ between 0.98-1.05.

Formal comparison between the different model variants (see above) was performed using the Deviance Information Criterion (Spiegelhalter et al., 2002), a commonly used method for assessing the goodness of fit in hierarchical models, for which a unique 'likelihood' is not defined, and the effective number of degrees of freedom is often unclear. Lower DIC values indicate a better fit, while taking into account the complexity of each model. A difference in DIC values of more than 10 is considered evidence for the winning model to capture the data significantly better.

MOTION ENERGY FILTERING AND PSYCHOPHYSICAL KERNELS

For the Visual motion 2AFC (FD) dataset, we used motion energy filtering (using the filters described in Urai and Wimmer, 2016) to reconstruct the time-course of fluctuating sensory evidence over the course of each individual trial, averaging over the spatial dimensions of the display (Supplementary Figure 7a,b). These single-trial traces then served as the time-resolved input to a set of extended DDM and leaky accumulator models (Figure 6). 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 (Kiani et al., 2008), yielding outputs that were a poor representation of the actual motion energy levels (see also Supplementary Figure 7a). 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 $T = \{1,2, \dots ,150\}$.

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 (Supplementary Figure 7c).

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 BIC 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. BIC is blind to such latent flexibility differences that do not map onto differences in number of parameters. Thus, BIC 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:

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

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

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

Above, g was the scaling parameter that controls the signal-to-noise-ration (given that c is fixed at 1). The variable $\ddot{\mathrm{v}}$ 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:

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

The variable $prev$ here encoded the previous choice (1: right, -1: left). If z_{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:

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

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:

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

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:

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

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:

$$\mathrm{dy(t)} = \mathrm{g} \cdot M_t \cdot \mathrm{dt} + \mathrm{cdW} \tag{12}$$

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:

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

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:

$$\mathrm{b}_{up}(t) = \left|\mathrm{a} - \mathrm{a}\frac{\mathrm{t}}{t+c}\right|_{\mathrm{a}/2}^{\mathrm{a}} \tag{14.1}$$

$$\mathrm{b}_{down}(t) = \left|\mathrm{a}\frac{\mathrm{t}}{t+c}\right|_{0}^{\mathrm{a}/2} \tag{14.2}$$

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:

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

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 (Busemeyer and Townsend, 1993; Zhang and Bogacz, 2010), constituting an extension of the DDM:

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

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:

$$\mathrm{dy(t)} = (\lambda \cdot \mathrm{y(t)} + \mathrm{g} \cdot M_t) \cdot \mathrm{dt} + \mathrm{cdW} \tag{17.1}$$

After stimulus offset, accumulation continued according to:

$$\mathrm{dy(t)} = \lambda \cdot \mathrm{y(t)} + \frac{y(T)}{T} + \mathrm{cdW} \tag{17.2}$$

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:

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

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 $\mathrm{y} = y_A - y_B$, with y_A and y_B being two independent accumulators coding the right and left choice. The two accumulators were initialized as follows:

$$y_A(0) = U(-sz, sz) \tag{19.1}$$

$$y_B(0) = 0 \tag{19.2}$$

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:

$$\mathrm{d}y_A(\mathrm{t}) = \left(\lambda_A \cdot y_A(\mathrm{t}) + \mathrm{g} \cdot f_A(M_t)\right) \cdot \mathrm{dt} + \frac{cdW}{\sqrt{2}} \tag{20.1}$$

$$\mathrm{d}y_B(\mathrm{t}) = \left(\lambda_B \cdot y_B(\mathrm{t}) + \mathrm{g} \cdot f_B(M_t)\right) \cdot \mathrm{dt} + \frac{cdW}{\sqrt{2}} \tag{20.2}$$

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:

$$f_A(x) = \begin{cases} x, if\ x > 0 \\ 0, if\ x \leq 0 \end{cases} \tag{20.3}$$

$$f_B(x) = \begin{cases} 0, if\ x > 0 \\ -x, if\ x \leq 0 \end{cases} \tag{20.4}$$

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

$$\lambda_A = \lambda + f_A(prev) \cdot \lambda_{bias} \tag{20.5}$$

$$\lambda_B = \lambda + f_B(prev) \cdot \lambda_{bias} \tag{20.6}$$

Leaky Accumulator with Static Bounds. A response was initiated when the difference between the two accumulators (*y*) exceeded a positive threshold $+\mathrm{a}$ (right choice) or surpassed a negative threshold $-\mathrm{a}$ (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 (*g*), 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 $\lambda - bias$ 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 $\mathrm{y} = y_A - y_B$ decision variable was at 0, rather than at a/2 as in all other models involving a single accumulator):

$$\mathrm{b}_{up}(t) = \left|\mathrm{a} - a\frac{\mathrm{t}}{t+c}\right|_0^{\mathrm{a}} \tag{21.1}$$

$$\mathrm{b}_{down}(t) = \left|\mathrm{a}\frac{\mathrm{t}}{t+c} - \mathrm{a}\right|_{-\mathrm{a}}^{0} \tag{21.2}$$

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:

$$L = \prod_k P_k^{n_k} \tag{23}$$

We applied a commonly used multi-stage approach to fit our simulation-based models (e.g. Teodorescu et al., 2016). 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*, BIC values were calculated at the group level:

$$BIC_f = -2\sum_{s}^{N} \ln(\mathcal{L}_s) + Nm_f \sum_{s}^{N} \ln(k_s) \tag{24}$$

where *N* is the total number of participants and *s* is the participants index. $\mathcal{L}_s$ and k_s denote the maximum likelihood estimate and number of trials 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 timesteps of 10 ms using equations 2 (DDM) and 18 (leaky accumulator).

For the DDM simulations, 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; Fig. Supplementary Figure 8a), three levels of constant drift bias (0.2, 0.5 and 0.8; Supplementary Figure 8b), three levels of a time-dependent linear increase in drift bias (1.5/s, 2.5/s and 3.5/s; Supplementary Figure 8c), 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 $c = 3$; Supplementary Figure 8d), 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, 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 (Supplementary Figure 7). We simulated three levels of starting point bias (0.05, 0.10 and 0.15; Supplementary Figure 8f), three levels of input bias (0.2, 0.5 and 0.8; Supplementary Figure 8g), and three

levels of λ-bias between the two accumulators: (-3 vs -2, -4 vs -1, and -5 vs 0; Supplementary Figure 8h).

We then fit DDM models separately to each of the three simulated leaky accumulator datasets (Supplementary Figure 8b, bottom). We fit the parameters boundary separation, drift rate, non-decision time, starting point, drift bias and drift rate variability.

EFFECTIVE BIAS SIGNAL

We calculated the effective bias signal (as in Hanks et al., 2011) for the winning DDM and leaky accumulator models with collapsing bounds (Figure 6d). We focused on those two models because they provided the best fits to the data overall (Figure 6b). 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.

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 (Katahira, 2016). The difference between two correlation coefficients that shared a common variable, and its associated p-value, was computed using Steiger's test (Steiger, 1980).

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) (Wetzels and Wagenmakers, 2012). We then multiplied these Bayes factors across datasets to quantify the total evidence in favor or against the null hypothesis of no correlation (Scheibehenne et al., 2016). 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 (Kass and Raftery, 1995).

DATA AND CODE AVAILABILITY

All behavioral data, model fits and analysis code is available at 10.6084/m9.figshare.7268558.

AUTHOR CONTRIBUTIONS

AEU, Conceptualization, Investigation, Formal Analysis, Software, Visualization, Writing – Original Draft, Writing – Review and Editing; JWdG, Investigation, Formal Analysis, Writing – Review and Editing; KT, Formal Analysis, Writing – Review and Editing; THD, Conceptualization, Resources, Supervision, Writing – Original Draft, Writing – Review and Editing.

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.

BIBLIOGRAPHY

Abrahamyan A, Silva LL, Dakin SC, Carandini M, Gardner JL (2016) Adaptable history biases in human perceptual decisions. Proc Natl Acad Sci 113:E3548–E3557.

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2:284–299.

Afacan-Seref K, Steinemann NA, Blangero A, Kelly SP (2018) Dynamic Interplay of Value and Sensory Information in High-Speed Decision Making. Curr Biol 28:795-802.e6.

Akaishi R, Umeda K, Nagase A, Sakai K (2014) Autonomous Mechanism of Internal Choice Estimate Underlies Decision Inertia. Neuron 81:195–206.

Allefeld C, Soon CS, Bogler C, Heinzle J, Haynes J-D (2013) Sequential dependencies between trials in free choice tasks. arxiv.org.

Basso MA, Wurtz RH (1997) Modulation of neuronal activity by target uncertainty. Nature 389:66–69.

Bode S, Sewell DK, Lilburn S, Forte JD, Smith PL, Stahl J (2012) Predicting Perceptual Decision Biases from Early Brain Activity. J Neurosci 32:12488–12498.

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

Bonaiuto JJ, Berker A de, Bestmann S (2016) Response repetition biases in human perceptual decisions are explained by activity decay in competitive attractor models. eLife 5:e20047.

Braun A, Urai AE, Donner TH (2018) Adaptive history biases result from confidence-weighted accumulation of past choices. J Neurosci 38:2418 –2429.

Brody CD, Hanks TD (2016) Neural underpinnings of the evidence accumulator. Curr Opin Neurobiol 37:149–157.

Busemeyer JR, Townsend JT (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. Psychol Rev 100:432.

Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Schölvinck ML, Zaharia AD, Carandini M (2011) The Detection of Visual Contrast in the Behaving Mouse. J Neurosci 31:11351–11361.

Chamberlain SR, Hampshire A, Müller U, Rubia K, del Campo N, Craig K, Regenthal R, Suckling J, Roiser JP, Grant JE, Bullmore ET, Robbins TW, Sahakian BJ (2009) Atomoxetine Modulates Right Inferior Frontal Activation During Inhibitory Control: A Pharmacological Functional Magnetic Resonance Imaging Study. Biol Psychiatry 65:550–555.

Chaudhuri R, Knoblauch K, Gariel M-A, Kennedy H, Wang X-J (2015) A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. Neuron 88:419–431.

Cheadle S, Wyart V, Tsetsos K, Myers N, de Gardelle V, Herce Castañón S, Summerfield C (2014) Adaptive Gain Control during Human Perceptual Choice. Neuron 81:1429–1441.

Cho RY, Nystrom LE, Brown ET, Jones AD, Braver TS, Holmes PJ, Cohen JD (2002) Mechanisms underlying dependencies of performance on stimulus history in a two-alternative forced-choice task. Cogn Affect Behav Neurosci 2:283–299.

Churchland AK, Kiani R, Shadlen MN (2008) Decision-making with multiple alternatives. Nat Neurosci 11:693–702.

Cisek P, Puskas GA, El-Murr S (2009) Decisions in Changing Conditions: The Urgency-Gating Model. J Neurosci 29:11560–11571.

de Gee JW, Colizoli O, Kloosterman NA, Knapen T, Nieuwenhuis S, Donner TH (2017) Dynamic modulation of decision biases by brainstem arousal systems. eLife 6:e23232.

de Gee JW, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. Proc Natl Acad Sci 111:E618–E625.

de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus Oscillatory Activity over Motor Cortex Reflects Perceptual Expectations. J Neurosci 33:1400–1410.

Drugowitsch J, Pouget A (2018) Learning optimal decisions with confidence. bioRxiv:244269.

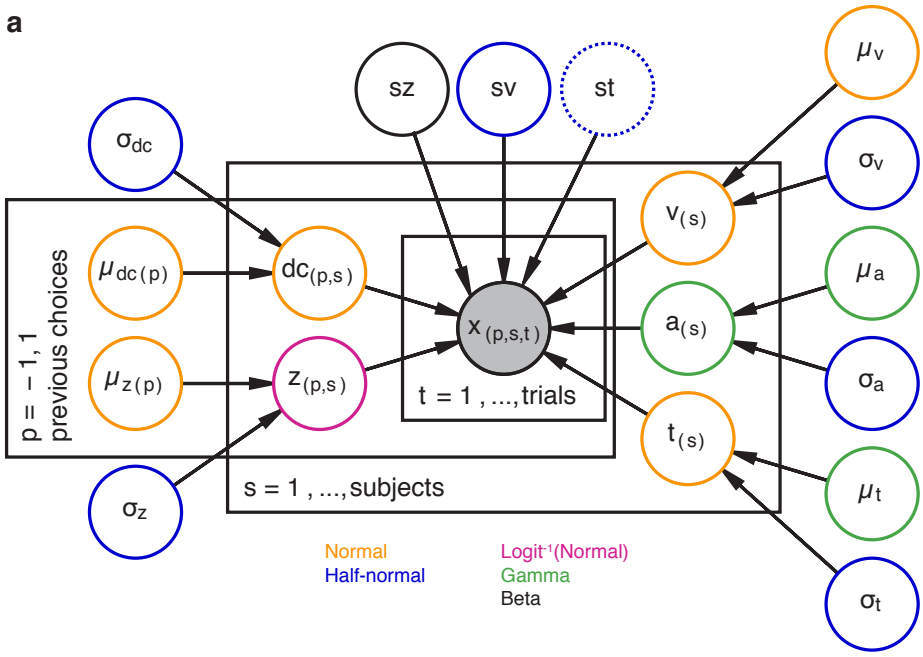
Dutilh G, van Ravenzwaaij D, Nieuwenhuis S, van der Maas HL, Forstmann BU, Wagenmakers E-J (2012a) How to measure post-error slowing: a confound and a simple solution. J Math Psychol 56:208–216.
Dutilh G, Vandekerckhove J, Forstmann BU, Keuleers E, Brysbaert M, Wagenmakers E-J (2012b) Testing theories of post-error slowing. Atten Percept Psychophys 74:454–465.
Edwards W (1965) Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. J Math Psychol 2:312–329.
Erlich JC, Brunton BW, Duan CA, Hanks TD, Brody CD (2015) Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. eLife 4.
Fernberger SW (1920) Interdependence of judgments within the series for the method of constant stimuli. J Exp Psychol 3:126.
Fründ I, Wichmann FA, Macke JH (2014) Quantifying the effect of intertrial dependence on perceptual decisions. J Vis 14:9–9.
Gao J, Tortell R, McClelland JL (2011) Dynamic Integration of Reward and Stimulus Information in Perceptual Decision-Making. PLoS ONE 6:e16749.
Gao J, Wong-Lin K, Holmes P, Simen P, Cohen JD (2009) Sequential Effects in Two-Choice Reaction Time Tasks: Decomposition and Synthesis of Mechanisms. Neural Comput 21:2407–2436.
Glaze CM, Kable JW, Gold JI (2015) Normative evidence accumulation in unpredictable environments. eLife 4:e08825.
Glimcher PW (2005) Indeterminacy in Brain and Behavior. Annu Rev Psychol 56:25–56.
Gold JI, Law C-T, Connolly P, Bennur S (2008) The Relative Influences of Priors and Sensory Evidence on an Oculomotor Decision Variable During Perceptual Learning. J Neurophysiol 100:2653–2668.
Gold JI, Shadlen MN (2007) The Neural Basis of Decision Making. Annu Rev Neurosci 30:535–574.
Goldfarb S, Wong-Lin K, Schwemmer M, Leonard NE, Holmes P (2012) Can post-error dynamics explain sequential reaction time patterns? Cogn Sci 3:213.
Green DM, Swets JA (1966) Signal Detection Theory and Psychophysics. John Wiley and Sons.
Hanks T, Kiani R, Shadlen MN (2014) A neural mechanism of speed-accuracy tradeoff in macaque area LIP. eLife 3:e02260.
Hanks TD, Mazurek ME, Kiani R, Hopp E, Shadlen MN (2011) Elapsed decision time affects the weighting of prior probability in a perceptual decision task. J Neurosci 31:6339–6352.
Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. Neuron 76:423–434.
Hwang EJ, Dahlen JE, Mukundan M, Komiyama T (2017) History-based action selection bias in posterior parietal cortex. Nat Commun 8:1242.
Jahfari S, Verbruggen F, Frank MJ, Waldorp LJ, Colzato L, Ridderinkhof KR, Forstmann BU (2012) How preparation changes the need for top-down control of the basal ganglia when inhibiting premature actions. J Neurosci Off J Soc Neurosci 32:10870–10878.
Jones M, Curran T, Mozer MC, Wilder MH (2013) Sequential effects in response time reveal learning mechanisms and event representations. Psychol Rev 120:628.
Kass RE, Raftery AE (1995) Bayes Factors. J Am Stat Assoc 90:773–795.
Katahira K (2016) How hierarchical models improve point estimates of model parameters at the individual level. J Math Psychol 73:37–58.
Kiani R, Hanks TD, Shadlen MN (2008) Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. J Neurosci 28:3017–3029.
Kim TD, Kabir M, Gold JI (2017) Coupled Decision Processes Update and Maintain Saccadic Priors in a Dynamic Environment. J Neurosci 37:3632–3645.
Leite FP, Ratcliff R (2011) What cognitive processes drive response biases? A diffusion model analysis. Judgm Decis Mak 6:651–687.
Leopold DA, Wilke M, Maier A, Logothetis NK (2002) Stable perception of visually ambiguous patterns. Nat Neurosci 5:605–609.
Liston DB, Stone LS (2008) Effects of prior information and reward on oculomotor and perceptual choices. J Neurosci 28:13866–13875.
Mazurek ME (2003) A Role for Neural Integrators in Perceptual Decision Making. Cereb Cortex 13:1257–1269.

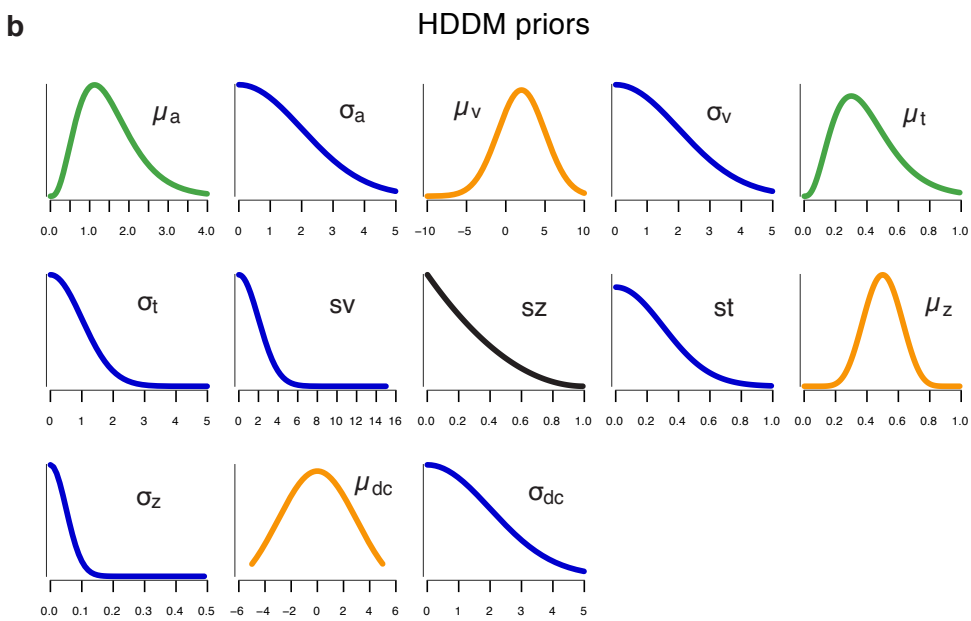
McGinley MJ, David SV, McCormick DA (2015) Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. Neuron 87:179–192.
Moran R (2015) Optimal decision making in heterogeneous and biased environments. Psychon Bull Rev 22:38–53.
Mulder MJ, Wagenmakers E-J, Ratcliff R, Boekel W, Forstmann BU (2012) Bias in the Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff. J Neurosci 32:2335–2343.
Murphy PR, Boonstra E, Nieuwenhuis S (2016) Global gain modulation generates time-dependent urgency during perceptual choice in humans. Nat Commun 7:13526.
Murphy PR, Vandekerckhove J, Nieuwenhuis S (2014) Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. PLoS Comput Biol 10:e1003854.
Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang X-J (2014) A hierarchy of intrinsic timescales across primate cortex. Nat Neurosci 17:1661–1663.
Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature 459:89–92.
Noorbaloochi S, Sharon D, McClelland JL (2015) Payoff Information Biases a Fast Guess Process in Perceptual Decision Making under Deadline Pressure: Evidence from Behavior, Evoked Potentials, and Quantitative Model Comparison. J Neurosci 35:10989–11011.
O'Connell RG, Shadlen MN, Wong-Lin K, Kelly SP (2018) Bridging Neural and Computational Viewpoints on Perceptual Decision-Making. Trends Neurosci 0 Available at: https://www.cell.com/trends/neurosciences/abstract/S0166-2236(18)30166-8 [Accessed July 14, 2018].
Ossmy O, Moran R, Pfeffer T, Tsetsos K, Usher M, Donner TH (2013) The Timescale of Perceptual Evidence Integration Can Be Adapted to the Environment. Curr Biol 23:981–986.
Padoa-Schioppa C (2013) Neuronal Origins of Choice Variability in Economic Decisions. Neuron 80:1322–1336.
Palminteri S, Wyart V, Koechlin E (2017) The Importance of Falsification in Computational Cognitive Modeling. Trends Cogn Sci 21:425–433.
Pape A-A, Siegel M (2016) Motor cortex activity predicts response alternation during sensorimotor decisions. Nat Commun 7:13098.
Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. Nature 400:233–238.
Purcell BA, Kiani R (2016a) Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex. Neuron 89:658–671.
Purcell BA, Kiani R (2016b) Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. Proc Natl Acad Sci:201524685.
Ratcliff R (2006) Modeling response signal and response time data. Cognit Psychol 53:195–237.
Ratcliff R, Childers R (2015) Individual Differences and Fitting Methods for the Two-Choice Diffusion Model of Decision Making. Decision 2:237–279.
Ratcliff R, McKoon G (2008) The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput 20:873–922.
Ratcliff R, Tuerlinckx F (2002) Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. Psychon Bull Rev 9:438–481.
Renart A, Machens CK (2014) Variability in neural activity and behavior. Curr Opin Neurobiol 25:211–220.
Reynolds JH, Heeger DJ (2009) The normalization model of attention. Neuron 61:168–185.
Rogers SL, Friedhoff LT (1998) Pharmacokinetic and pharmacodynamic profile of donepezil HCl following single oral doses. Br J Clin Pharmacol 46.
Rokem A, Silver MA (2010) Cholinergic Enhancement Augments Magnitude and Specificity of Visual Perceptual Learning in Healthy Humans. Curr Biol 20:1723–1728.
Rorie AE, Gao J, McClelland JL, Newsome WT (2010) Integration of Sensory and Reward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP) of the Macaque Monkey. PLOS ONE 5:e9308.
Roxin A, Ledberg A (2008) Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. PLoS Comput Biol 4:e1000046.
Runyan CA, Piasini E, Panzeri S, Harvey CD (2017) Distinct timescales of population coding across cortex. Nature 548:92–96.

Sauer J-M, Ring BJ, Witcher JW (2005) Clinical pharmacokinetics of atomoxetine. Clin Pharmacokinet 44:571–590.
Scheibehenne B, Jamil T, Wagenmakers E-J (2016) Bayesian Evidence Synthesis Can Reconcile Seemingly Inconsistent Results: The Case of Hotel Towel Reuse. Psychol Sci 27:1043–1046.
Scott BB, Constantinople CM, Akrami A, Hanks TD, Brody CD, Tank DW (2017) Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. Neuron 95:385–398.
Servant M, Montagnini A, Burle B (2014) Conflict tasks and the diffusion framework: Insight in model constraints based on psychological laws. Cognit Psychol 72:162–195.
Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci 16:1486–1510.
Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol 64:583–639.
St. John-Saaltink E, Kok P, Lau HC, de Lange FP (2016) Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. J Neurosci 36:6186–6192.
Stanislaw H, Todorov N (1999) Calculation of signal detection theory measures. Behav Res Methods Instrum Comput 31:137–149.
Steiger JH (1980) Tests for comparing elements of a correlation matrix. Psychol Bull 87:245.
Teodorescu AR, Moran R, Usher M (2016) Absolutely relative or relatively absolute: violations of value invariance in human decision making. Psychon Bull Rev 23:22–38.
Thura D, Guberman G, Cisek P (2016) Trial-to-trial adjustments of speed-accuracy trade-offs in premotor and primary motor cortex. J Neurophysiol 117:665–683.
Treisman M, Williams TC (1984) A theory of criterion setting with an application to sequential dependencies. Psychol Rev 91:68.
Tsetsos K, Chater N, Usher M (2012) Salience driven value integration explains decision biases and preference reversal. Proc Natl Acad Sci 109:9659–9664.
Tsetsos K, Pfeffer T, Jentgens P, Donner TH (2015) Action Planning and the Timescale of Evidence Accumulation. PLoS ONE 10:e0129473.
Ulrich R, Schröter H, Leuthold H, Birngruber T (2015) Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. Cognit Psychol 78:148–174.
Urai AE, Braun A, Donner TH (2017) Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat Commun 8:14637.
Urai AE, Wimmer K (2016) Spatiotemporal motion energy filtering: a Matlab implementation. Zenodo. Available at: 10.5281/zenodo.167351.
Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. Psychol Rev 108:550.
van Ravenzwaaij D, Mulder MJ, Tuerlinckx F, Wagenmakers E-J (2012) Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. Front Psychol 3:132.
Voss A, Rothermund K, Brandtstädter J (2008) Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. J Exp Soc Psychol 44:1048–1056.
Wei W, Wang X-J (2016) Inhibitory Control in the Cortico-Basal Ganglia-Thalamocortical Loop: Complex Regulation and Interplay with Memory and Decision Processes. Neuron 92:1093–1105.
Wetzels R, Wagenmakers E-J (2012) A default Bayesian hypothesis test for correlations and partial correlations. Psychon Bull Rev 19:1057–1064.
White CN, Poldrack RA (2014) Decomposing bias in different types of simple decisions. J Exp Psychol Learn Mem Cogn 40:385–398.
Wiecki TV, Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. Front Neuroinformatics 7.
Wilder M, Jones M, Mozer MC (2009) Sequential effects reflect parallel learning of multiple environmental regularities. In: Advances in Neural Information Processing Systems 22 (Bengio Y, Schuurmans D, Lafferty JD, Williams CKI, Culotta A, eds), pp 2053–2061. Curran Associates, Inc. Available at: http://papers.nips.cc/paper/3870-sequential-effects-reflect-parallel-learning-of-multiple-environmental-regularities.pdf [Accessed November 10, 2015].

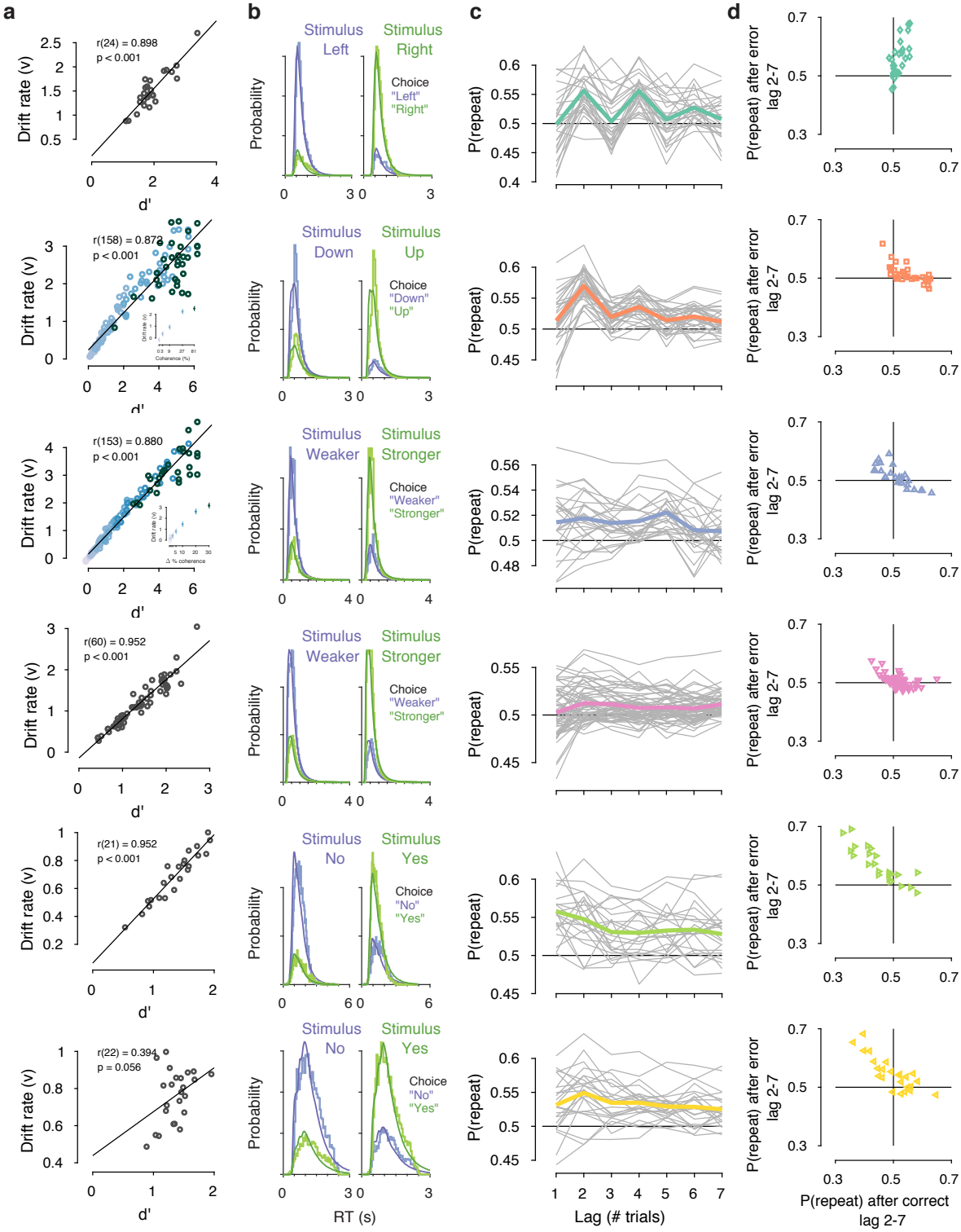
Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J (2015) Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat Commun 6:6177.
Wong K-F, Wang X-J (2006) A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. J Neurosci 26:1314–1328.
Wyart V, Koechlin E (2016) Choice variability and suboptimality in uncertain environments. Curr Opin Behav Sci 11:109–115.
Wyart V, Nobre AC, Summerfield C (2012) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. Proc Natl Acad Sci 109:3593–3598.
Yu AJ, Cohen JD (2008) Sequential effects: Superstition or rational behavior? Adv Neural Inf Process Syst 21:1873–1880.
Zhang J, Bogacz R (2010) Bounded Ornstein–Uhlenbeck models for two-choice time controlled tasks. J Math Psychol 54:322–333.
Zhang S, Huang CH, Yu AJ (2014) Sequential effects: a Bayesian analysis of prior bias on reaction time and behavioral choice. In: Proceedings of the 36th Annual Conference of the Cognitive Science Society Available at: https://mindmodeling.org/cogsci2014/papers/320/paper320.pdf [Accessed August 5, 2016].

SUPPLEMENTARY FIGURES

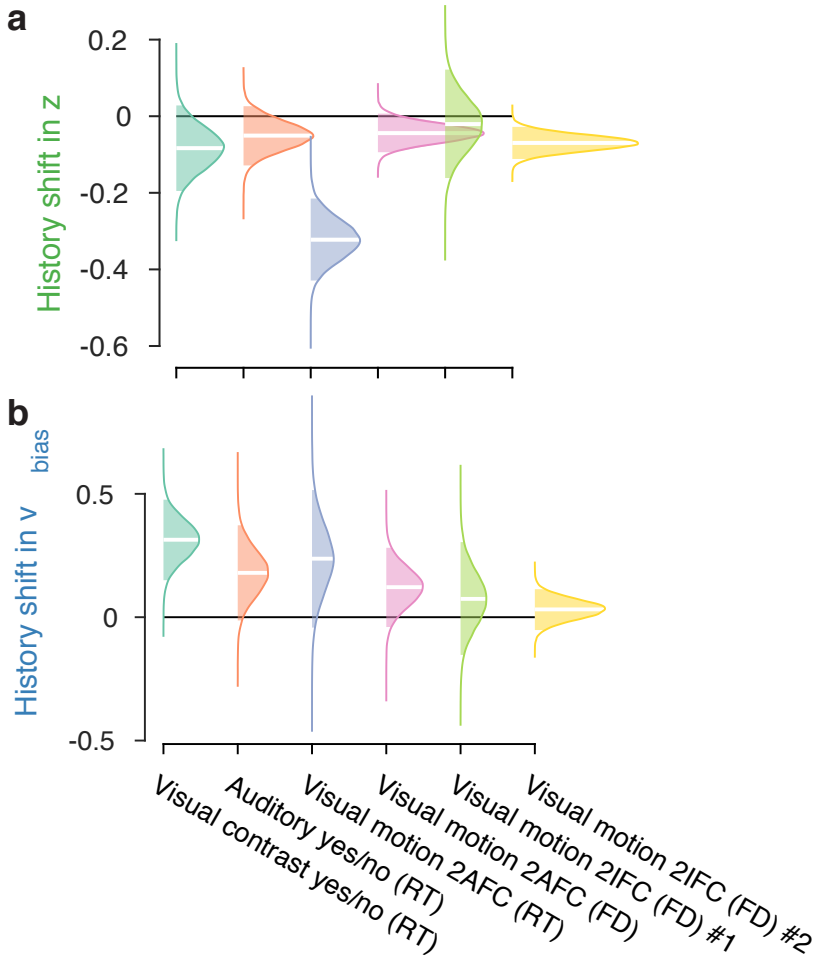




Supplementary Figure 1. The hierarchical DDM. *(**a**) Graphical representation of the hierarchical model structure. The full model (with both history-dependent drift bias and starting point) is depicted. Round nodes represent random variables, and the shaded node x represented the observed data (choices 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 (Supplementary Figure 2a, inset). Between-trial variability in non-decision time was only included in the model shown in Supplementary Figure 3b. (**b**) Prior distributions used for each node, with colors indicating their distribution. See (Wiecki et al., 2013) for the full set of prior specifications.*

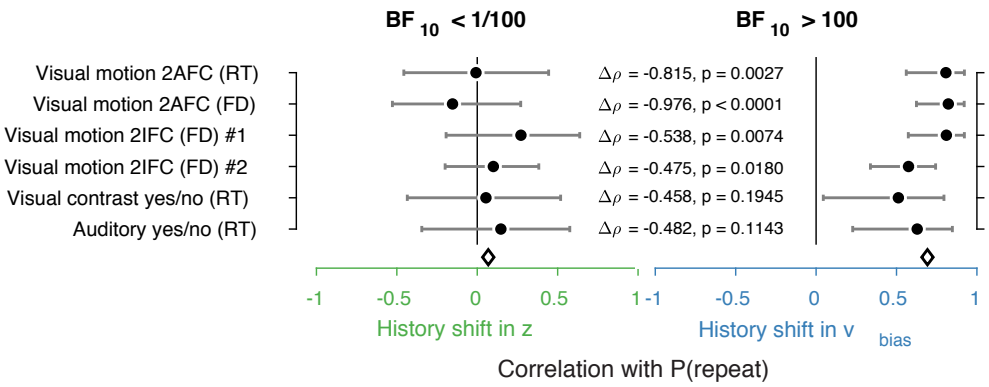

a
b
c
d
Drift rate (v)
d'
r(24) = 0.898
p < 0.001
r(158) = 0.873
p < 0.001
r(153) = 0.880
p < 0.001
r(60) = 0.952
p < 0.001
r(21) = 0.952
p < 0.001
r(22) = 0.394
p = 0.056
Probability
Stimulus Left
Stimulus Right
Choice "Left" "Right"
Stimulus Down
Stimulus Up
Choice "Down" "Up"
Stimulus Weaker
Stimulus Stronger
Choice "Weaker" "Stronger"
Stimulus No
Stimulus Yes
Choice "No" "Yes"
RT (s)
P(repeat)
Lag (# trials)
P(repeat) after error lag 2-7
P(repeat) after correct lag 2-7

Supplementary Figure 2. Drift diffusion model and choice history patterns. *(**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. (**c**) Probability of repeating previous choices, for each of 1-7 trials back. Thick, colored line indicates the average over participants. (**d**) Repetition probability after correct vs. error trials, averaged over lag 2-7 (complementary to Figure 2c).*

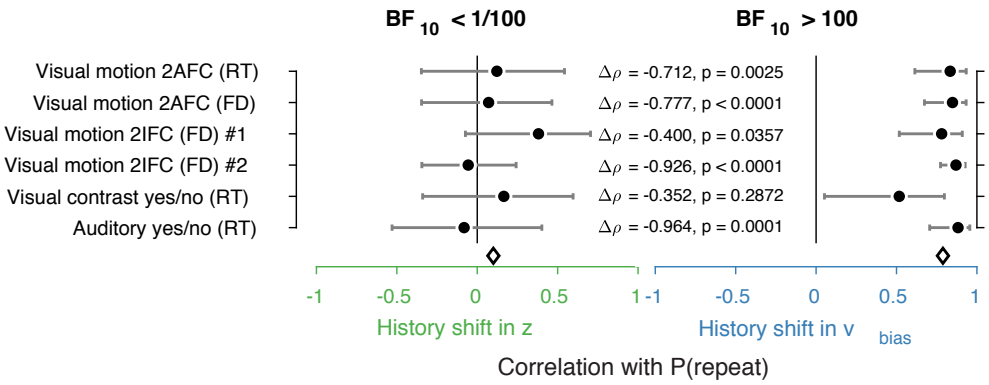


Supplementary Figure 3. Group-level posterior distributions of history bias parameters. *History shift in (**a**) starting point and (**b**) drift bias, separately for each dataset. The posteriors were taken from the model where history-dependent z and* v_{bias} *were fit simultaneously. Distributions were smoothed using kernel density fits. Shaded regions represent the 95% BCI, and white lines indicate the posterior mean.*

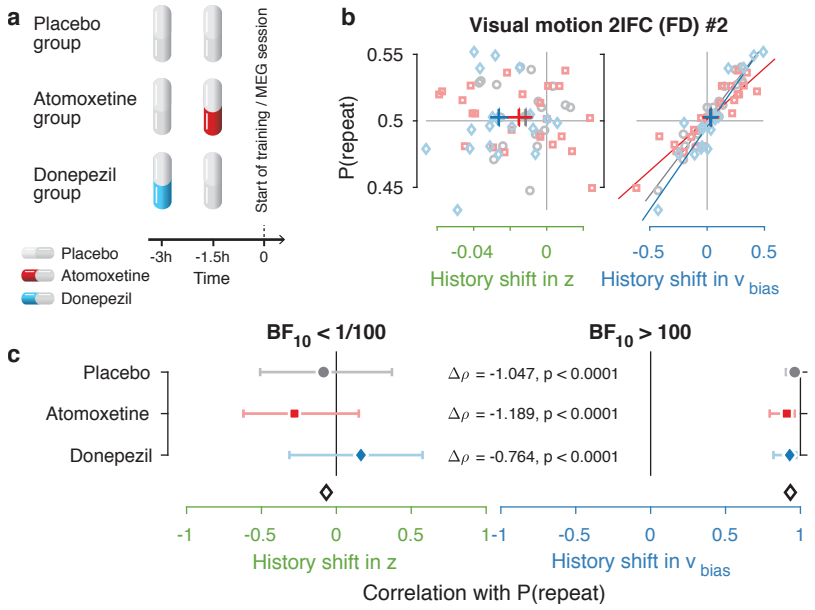
a. Non-hierarchical G^2 quantile fit



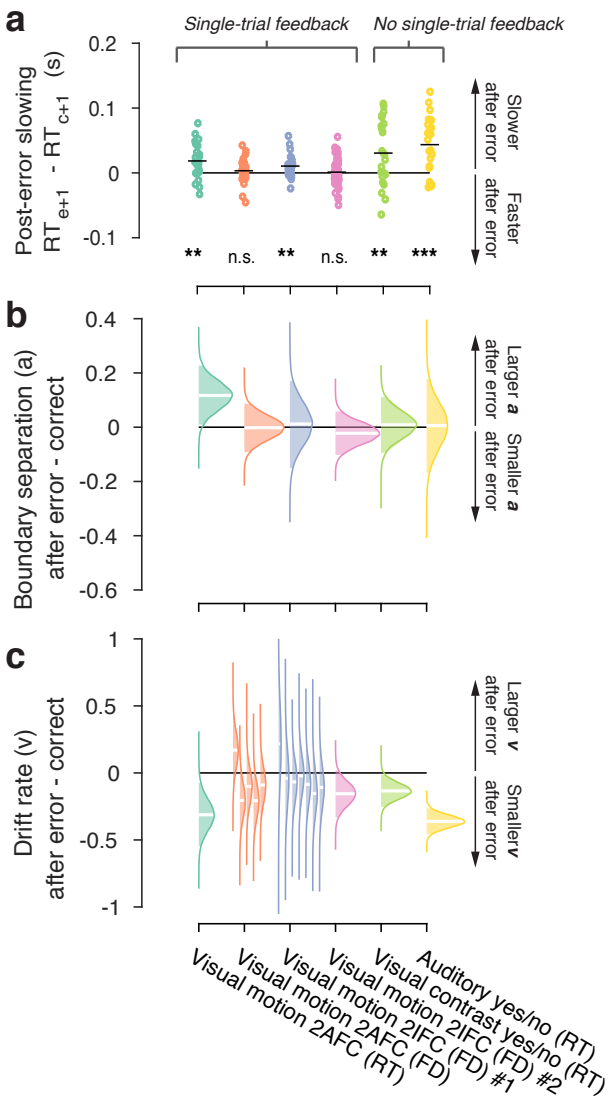
b. Non-zero across-trial variability in drift rate (v), starting point (z), and non-decision time (t)



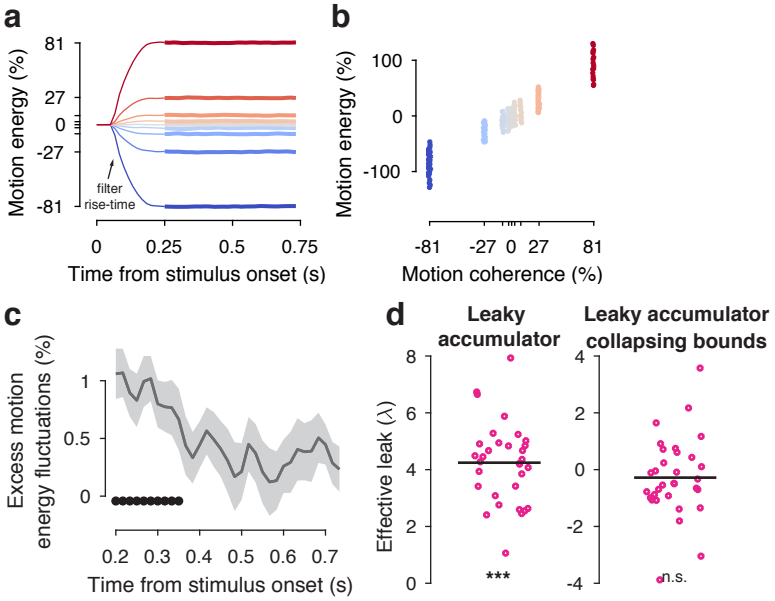
Supplementary Figure 4. Control model fits. *(**a**) Summary figure based on non-hierarchical G^2 fits (Ratcliff and Tuerlinckx, 2002). Rather than the full RT distribution, we fit each observers' RT quantiles (0.1, 0.3, 0.5, 0.7, 0.9) and correlated history shifts in the DDM bias parameters to individual 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 (Ratcliff and Childers, 2015).*



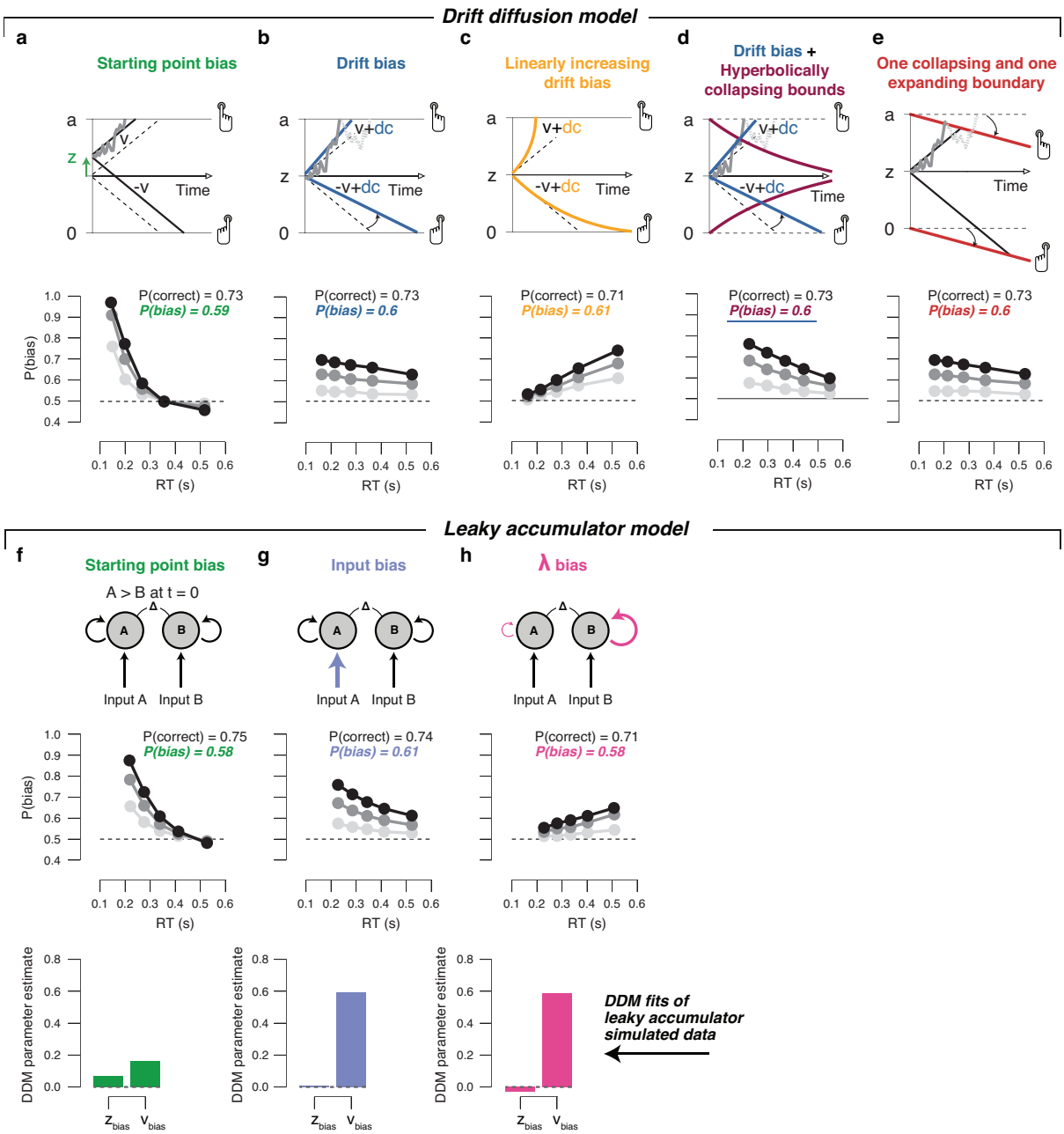
Supplementary Figure 5. Same biasing mechanism under two pharmacological interventions. *(**a**) Participants in the MEG study were assigned to one of three pharmacological groups. At the start of each experimental session, they orally took 40 mg atomoxetine (Strattera®), 5 mg donepezil (Aricept®), or placebo. Since the time of peak plasma concentration is 3 hours for donepezil (Rogers and Friedhoff, 1998) and 1-2 hours for atomoxetine (Sauer et al., 2005), 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 (Chamberlain et al., 2009; Rokem and Silver, 2010). 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.*



Supplementary Figure 6. 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. (**a**) Changes in boundary separation after error vs. correct trials. (**b**) 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 Supplementary Figure 2a). 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.*



Supplementary Figure 7. 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.*



Supplementary Figure 8. Extended model simulations. (a-e) *Simulated conditional bias functions (fraction of biased choices as a function of RT; 100K trials per model) for various biasing mechanisms within the framework of the DDM (see Methods).* ***(a)*** *The effect of a biased starting point on choice declines rapidly with elapsed time, as the decision variable is increasingly governed by the drift rate.* ***(b)*** *The effect of a constant drift bias on choice stays relatively constant within the trial, decreasing but remaining positive also for slow RTs.* ***(c)*** *The effect of a linearly increasing drift bias on choice increases with elapsed time, as the decision variable is increasingly governed by the drift bias.* ***(d)*** *The effect of a drift bias combined with hyperbolically collapsing bounds (Figure 6a; Hanks et al., 2011) closely resembles that of drift bias alone; differences include overall shorter and slightly less accurate decisions.* ***(e)*** *One time-dependent collapsing and one expanding bound are equivalent to time-dependent changes in the decision variable due to drift bias. These different biasing mechanisms*

produce identical changes in the conditional response functions. ***(f-h)*** *Simulated conditional bias functions (100K trials per model) for various biasing mechanisms within the framework of the leaky accumulator model with collapsing bounds, corresponding to Figure 6b5 (Materials and Methods). To match the empirical data (Supplementary Figure 7d), 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).* ***(f)*** *The effect of a biased starting point on choice declines rapidly with elapsed time, as the decision variable is increasingly governed by the input. Bottom: A starting point bias within the leaky accumulator model loads on both DDM starting point and drift bias within the standard DDM. This can be explained by the λ parameter, which determines that any starting point bias of the accumulators will have a longer lasting effect on the decision variable.* ***(g)*** *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.* ***(h)*** *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. Note the close correspondence with a linearly increasing DDM drift bias (**e**). Bottom: leaky accumulator λ bias is primarily explained by DDM drift bias; the DDM fits an additionally starting point bias in the opposite direction to push down the predicted fraction of biased choices for early RTs. All panels: Gray line, example trajectory of decision variable from single trial; dashed lines in the conditional response function plots, conditional response function for unbiased model; the numbers printed in the condition response function plots are overall probabilities, regardless of RT.*

*It noteworthy that λ bias in the leaky accumulator model (**h**) predicts the strongest choice bias for long RTs, while both DDM starting point and drift bias predict the strongest choice bias for short RTs (**a**, **b**). 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 (Supplementary Figure 3). 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.*