SUPPLEMENTARY INFORMATION

1. Supplementary MAP results

1.1. The Hybrid model underestimates perseveration in PSL agents

To provide further evidence for the tradeoff between perseveration and learning asymmetries reported in the main text (Figure 2c), we directly compared how the PSL and Hybrid models estimate the perseveration parameter from the same underlying data. This analysis used the same simulations of Perseverative Symmetrically-Learning (PSL) agents described in the main text, which were generated using the empirically-derived parameters from each participant across the 10 experiments.

We fitted these simulated datasets with two different models: the true generative model (PSL) and the Hybrid model. As shown in Figure S1, the Hybrid model fits consistently returned a lower estimate for the perseveration parameter (ϕ) than the PSL model fits (all p's<0.05; except for C4, p=0.71).

This figure demonstrates that the Hybrid model underestimates perseveration. When considered alongside Figure 2, it shows how the spurious confirmation bias emerges: the Hybrid model incorrectly partitions the variance from the PSL agent, misattributing a portion of the perseverative tendency (which it underestimates, as shown here) to a spurious learning asymmetry (which it overestimates, as shown in Figure 2). This direct comparison therefore provides a clear illustration of the mechanistic tradeoff that drives the model mimicry problem.

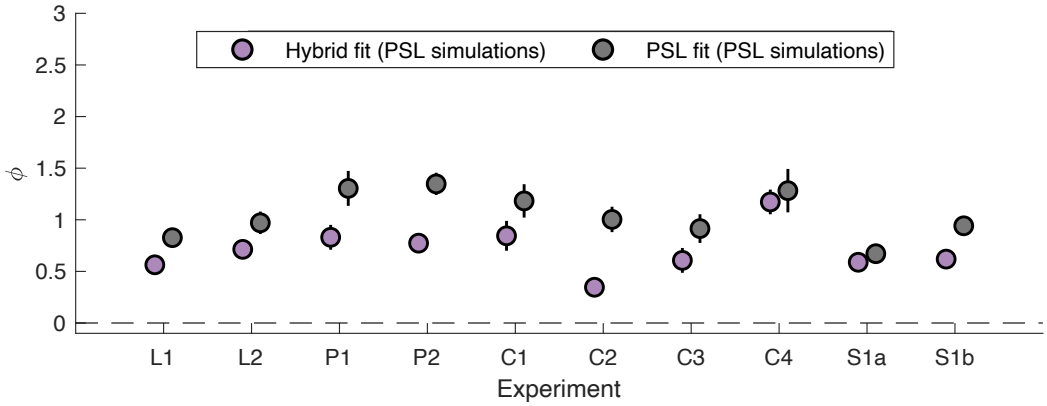

Figure S1: Comparison of perseveration parameter (phi) when fitting PSL model simulations with either PSL model or the Hybrid model. *The Hybrid model (violet) consistently yields a significantly lower estimate of phi than the PSL model (gray). Points and error bars represent the mean ± SEM across simulations for each experiment.*

1.2. Choice Perseveration Generates Spurious Confirmation Bias in Expanded Models

To test the generalizability of our findings, we repeated the analysis pipeline from Figure 2 using the more complex "expanded" models that have been proposed for full-feedback tasks. These models allow for separate learning rates for chosen (α_c^+, α_c^-) and unchosen options (α_u^+, α_u^-). Consequently, the Q-values of chosen options are updated as:

$$Q_{t+1}(C) = \begin{cases} Q_t(C) + \alpha_c^+ \cdot \delta_t(C) \; if \; \delta_t(C) > 0 \\ Q_t(C) + \alpha_c^- \cdot \delta_t(C) \; if \; \delta_t(C) < 0 \end{cases}$$

And unchosen options are updated as:

$$Q_{t+1}(C) = \begin{cases} Q_t(C) + \alpha_u^+ \cdot \delta_t(C) \; if \; \delta_t(C) > 0 \\ Q_t(C) + \alpha_u^- \cdot \delta_t(C) \; if \; \delta_t(C) < 0 \end{cases}$$

The Expanded CB model uses these four learning rates and the standard softmax rule. It has 5 free parameters: $\{\alpha_c^+, \alpha_c^-, \alpha_u^+, \alpha_u^-, \beta\}$. The Expanded Hybrid model uses these four learning rates and the hybrid softmax rule. It has 7 free parameters: $\{\alpha_c^+, \alpha_c^-, \alpha_u^+, \alpha_u^-, \beta, \tau, \phi\}$. Confirmation bias is defined as $(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/2)$ in absolute terms, and as $(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/(\alpha_c^+ + \alpha_u^- + \alpha_c^- + \alpha_u^+)$ in normalized terms.

Previous work by Sugawara and Katahira (2021) suggested a perfect dissociation between perseveration and confirmation bias for these models, though this was not tested for empirical parameters and across several task designs. Here, we repeat the analyses from our main results (Figure 2 in main text) for these expanded models, but restricted this analysis to the three experiments with full feedback (P2, C2, and S1b). As in our main analysis, we found that the detected confirmation bias (both absolute and normalized) was significantly reduced when moving from the Expanded CB model to the Expanded Hybrid model (absolute, all p's<0.05 except for P2, normalized, all p's<0.05). Crucially, when we simulated PSL agents (using the empirically-derived MAP parameters) and fitted them with the Expanded Hybrid model, we still found a significant spurious confirmation bias across all three tasks (all p's<0.001). This was accompanied by a significant underestimation of the perseveration parameter in the PSL simulations when fitted with the Expanded Hybrid model (p<0.001). These results demonstrate that the confound between perseveration and learning asymmetries persists even with four learning rates.

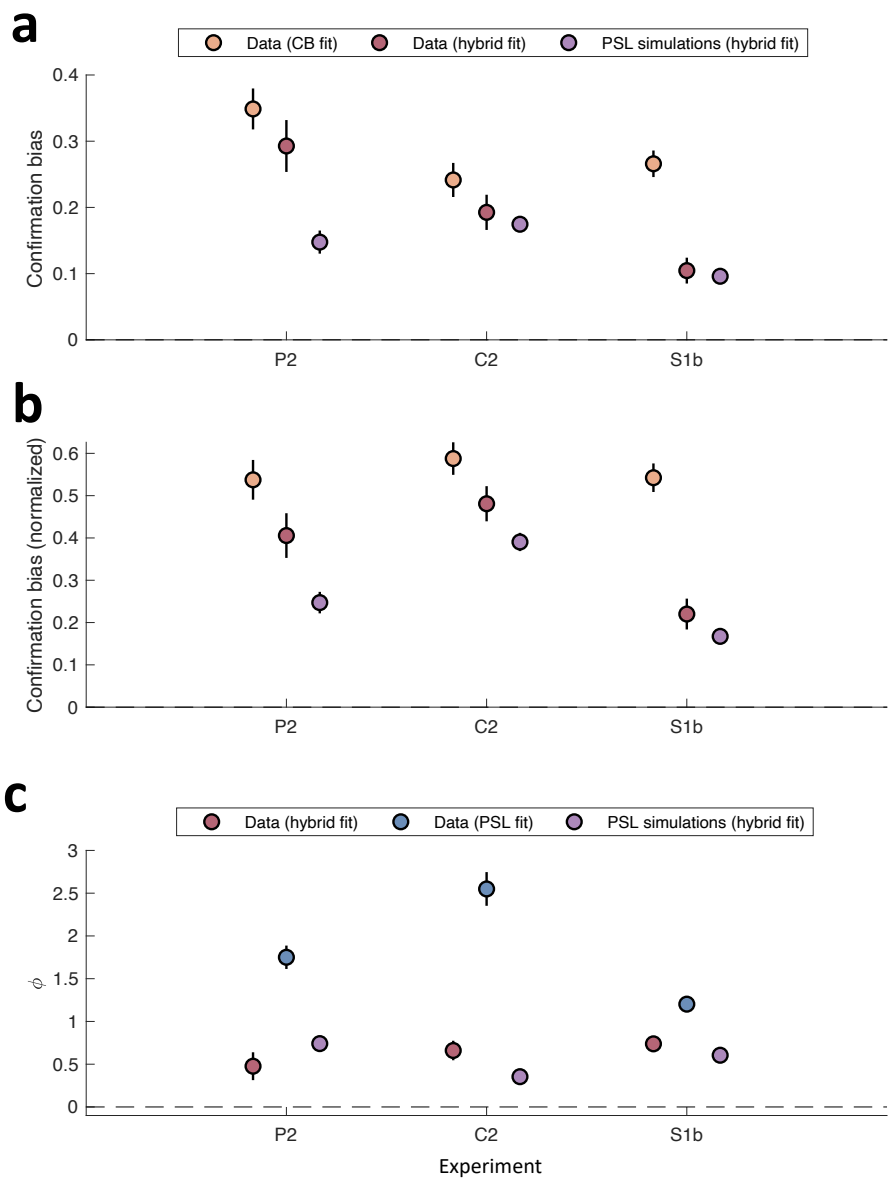

Figure S2: Confirmation bias and perseveration parameters estimated with Expanded Models. ***(a)*** *Absolute confirmation bias* $[(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/2)]$ *estimated from fitting participant data with the Expanded CB model (yellow points) and the Expanded Hybrid model (red points). Additionally, violet points show the spurious confirmation bias recovered when fitting PSL simulations with the Expanded Hybrid model.* ***(b)*** *Same as (a), but for the normalized confirmation bias. The normalized version of confirmation bias is defined as* $(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/(\alpha_c^+ + \alpha_u^- + \alpha_c^- + \alpha_u^+)$ ***(c)*** *The perseveration parameter (ϕ) estimated from participant data using the Expanded Hybrid model (red points) and the PSL model (blue points). Violet points show the ϕ values recovered from fitting the Hybrid model to simulations generated by the PSL model, whose original generative parameters are shown in blue. The recovered violet points are systematically lower than the generative blue points, indicating underestimation. Points and error bars represent mean ± SEM across participants/simulations for the three full-feedback experiments.*

1.3. MAP Estimation Introduces Spurious Asymmetries in the Absence of Perseveration

To further isolate the source of the parameter estimation biases, we conducted a control analysis to test whether the MAP fitting procedure could introduce a spurious confirmation bias even in the complete absence of perseveration.

To this end, we first created a simple symmetric learning model, which is a version of the CB model with a single, symmetric learning rate (α) and no perseveration parameters ($\phi = 0, \tau = 0$). We fitted this simple model to the empirical data from each participant to obtain their best-fit parameters under this basic learning assumption. Next, we used these empirically-derived parameters to generate 1,001 simulated datasets per participant. By construction, these datasets represented the behavior of agents who learn symmetrically and have no perseverative tendencies. Finally, we fitted these simulated datasets with the full Hybrid model using the MAP estimation procedure, and we averaged such fitted parameters across their corresponding 1,001 simulations.

Ideally, if the fitting procedure were unbiased, the recovered parameters for confirmation bias and perseveration should be centered around zero. However, our analysis revealed a systematic bias. We found a significant, spurious confirmation bias in several of the simulated experiments (absolute, $p<0.001$ in P1, P2, C2 and S1b; normalized, $p<0.05$ in all except C1,C4 and S1a; Fig. S3a-b). Furthermore, the perseveration parameter (ϕ) was consistently estimated to be significantly *below* zero across all experiments (all p's<0.001; Fig. S3c).

These results show that the MAP procedure itself is biased. A significant confirmation bias can be detected even when the underlying generative process contains neither perseveration nor learning asymmetries, highlighting a source of estimation bias that is independent of the model mimicry problem.

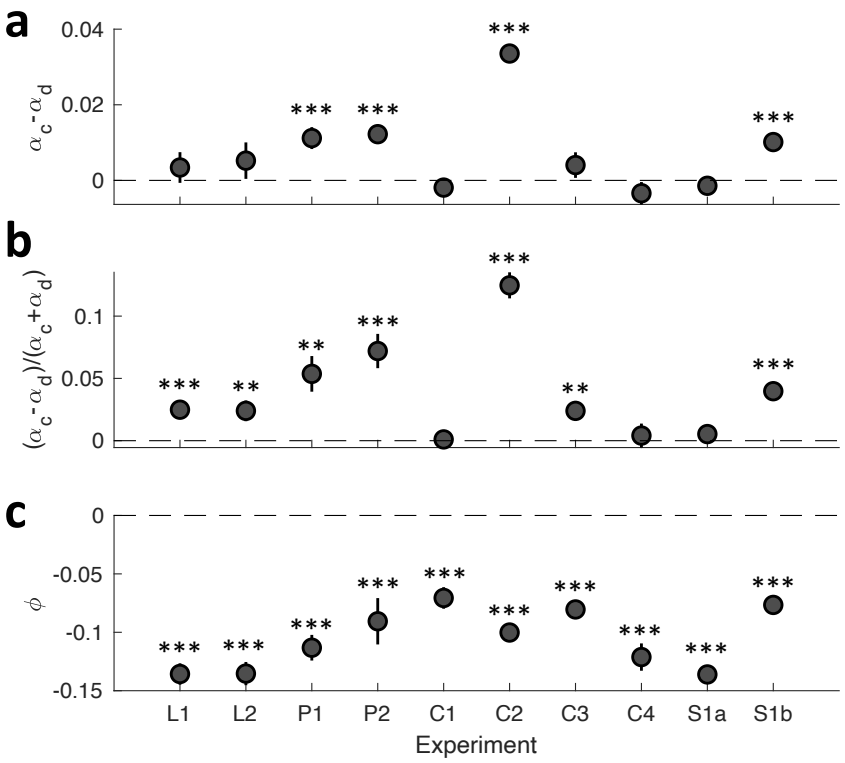

Figure S3: Spurious parameters recovered from simulations of a simple symmetric learning model. *The figure shows the results of fitting the Hybrid model to simulated data from a simple model with symmetric learning and no perseveration.* ***(a)*** *Absolute confirmation bias fitted from the simulations.* ***(b)*** *Same as (a), but for the normalized confirmation bias.* ***(c)*** *The perseveration parameter (ϕ) fitted from the simulations. Across all panels, the ideal unbiased estimate would be zero. Points and error bars represent the mean ± SEM across simulations. (* p < 0.05; ** p < 0.01; *** p < 0.001).*

1.4. Confirmation bias Does not Generate Spurious Positive Perseveration

Having shown that perseveration can be misidentified as a confirmation bias, we next tested whether the model mimicry we identified operates in the opposite direction. Specifically, we investigated whether behavior generated by a pure confirmation bias could lead to the spurious detection of perseveration. To this aim, we simulated 1001 datasets per participant from our CB model (using the MAP-estimated empirical parameters) and subsequently fitted these datasets with our hybrid model.

We found that the confirmation bias was accurately recovered. The fitted confirmation bias parameters were not significantly different from the generative empirical parameters in nearly all cases (yellow vs green points in Fig. S4a-b; absolute, all p's>0.05, except for S1b p<0.001; normalized, all p's>0.05). We did not find a spurious *positive* perseveration. Instead, the fitted perseveration parameters (ϕ) were significantly *negative* for most experiments (L1, L2, P1, C3, C4, S1a; all p's<0.05), non-significant for three experiments (P2, C1, C2), and only significantly positive for one experiment (S1b) (Fig. S4c). Together, these results suggest that the confound is largely unidirectional. While perseveration can be misattributed as a confirmation bias, a

confirmation bias is generally recovered accurately by the hybrid model and is not misattributed as a positive perseveration.

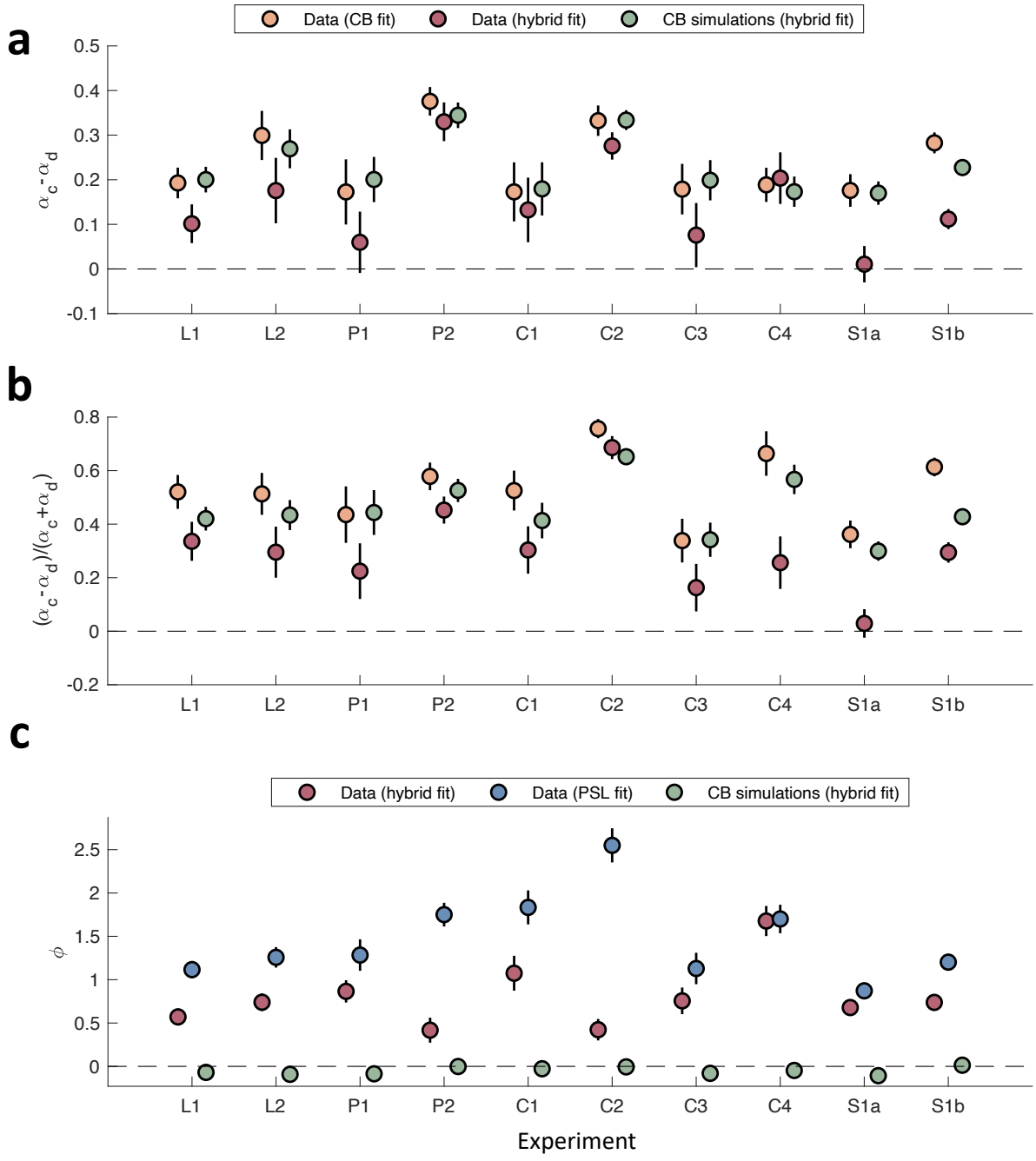

Figure S4: Confirmation bias and perseveration parameters estimated from empirical data and CB simulations. ***(a)*** *Absolute confirmation bias estimated from fitting participant data with the Pure CB model (yellow points) and the Hybrid model (red points). Additionally, green points show the confirmation bias recovered when fitting CB simulations with the Hybrid model.* ***(b)*** *Same as (a), but for the normalized confirmation bias.* ***(c)*** *The perseveration parameter (ϕ) estimated from participant data using the Hybrid*

model (red points) and the PSL model (blue points). Green points show the ϕ values recovered from fitting the Hybrid model to simulations of CB agents, showing that no spurious positive perseveration parameter emerges for most experiments. Points and error bars represent mean ± SEM across participants/simulations.

1.5. Asymptotic Dissociation of Perseveration and Confirmation Bias in Large Datasets

Here, we replicate the asymptotic analysis from the main text (Figure 4), but using the MAP estimation procedure to investigate whether the use of parameter priors exacerbates the small-sample bias. To do this, we first simulated PSL agents using the average empirical parameters obtained from fitting experiment P2 with the MAP procedure. We systematically varied the length of the experiments, creating datasets that included 2, 4, 8, 16, and so on, up to 1024 sessions. For each dataset length, we generated 200 independent simulations. We then fitted each of these simulated datasets with the Hybrid model using the MAP procedure and examined the estimation of all four key parameters – the confirmatory (α_c) and disconfirmatory (α_d) learning rates, and the perseveration parameters (ϕ and τ) – as a function of dataset size.

Our analysis shows that with MAP estimation, all parameters are misestimated in small datasets and require an unfeasibly large number of trials to converge to their true generative values. First, we found that the overall spurious confirmation bias only vanished after an extremely large number of trials – around 100,000 – were included in the simulation (Fig. S5a). This artifact was driven by a systematic overestimation of $\alpha_c$coupled with an even greater underestimation of α_d. Both learning rates exhibited extremely slow convergence, only approaching their true generative values in datasets of around 100,000 trials (Fig. S5b).

The perseveration parameters were also misestimated. The choice-trace accumulation rate (τ) was initially overestimated, converging to its true generative value in datasets of around 6,000 trials (Fig. S5c). In contrast, ϕ was severely underestimated. This is a direct consequence of the Gaussian prior used in the MAP procedure, which shrinks the parameter estimate towards zero. As a result, its explanatory power is incorrectly absorbed by the artifactual learning bias. This underestimation was so severe that it only converged to the true generative value in datasets with an extremely large number of trials (around 100,000 trials) (Fig. S5d).

These results highlight the profound impact of parameter priors on estimation. The strong shrinkage effect on the ϕ parameter requires a massive amount of data to overcome, explaining why the MAP procedure needs many more trials to achieve an unbiased dissociation of the mechanisms compared to the MLE procedure.

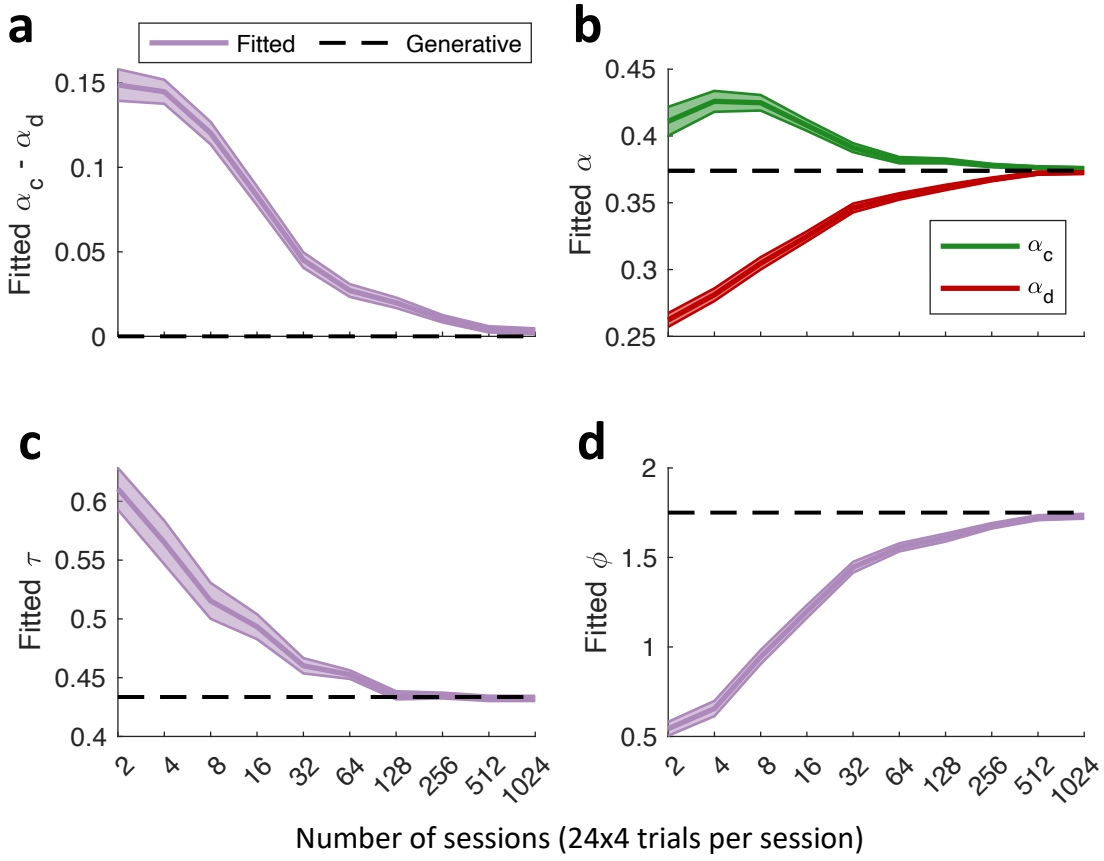

Figure S5: Misestimation of parameters fitted with MAP estimation is reduced with dataset size. *The figure shows parameter estimates obtained by fitting the Hybrid model to simulations of PSL agents, as a function of the number of sessions in the simulated dataset (where each session encompasses 4 blocks of 24 trials). The simulations were based on the task in the second experiment (full feedback condition) from Palminteri et al. (2017)).* ***(a)*** *The absolute confirmation bias ($\alpha_c - \alpha_d$) fitted with MLE is large for small datasets and asymptotes towards zero for large datasets (around 100,000 trials).* ***(b)*** *Disaggregation of the effect in (a), showing the separately recovered confirmatory (α_c) and disconfirmatory (α_d) learning rates.* ***(c)*** *Fitted perseveration parameter (ϕ). It is underestimated in small datasets and asymptotes towards its true value for datasets of around 100,000 trials. In all panels, solid lines and shaded areas represent the mean ± SEM across 200 simulations. Dashed lines represent the true generative parameter value used in the simulations. Gray boxes highlight dataset sizes for which the recovered parameter is significantly different from the generative value.*

1.6. Contribution of different parameters to the spurious confirmation bias

To better understand the mechanics of the artifact, we conducted a parameter sweep analysis to characterize which parameters of the PSL model most strongly drive the emergence of the artifactual confirmation bias. To do this, we systematically varied each of the four key generative parameters of the PSL model (β, α, τ, and ϕ) across a range of values. For each step of the sweep, one parameter was varied across a grid of values, while the other three parameters were held fixed at the mean of the best-fit estimates from the original PSL MAP fits of experiment P2.

For each combination, we generated 1,001 datasets for the P2 task, fitted them with the full hybrid model, and examined the resulting confirmation bias.

This analysis revealed complex relationships between the generative parameters and the artifactual bias. First, the largest bias emerged when the value-based choice inverse temperature (β) was low; this effect was driven by an overestimation of the confirmatory learning rate (α_c) and an even greater underestimation of the disconfirmatory learning rate (α_d) (Fig. S6a-c, left panels). The true learning rate (α) had more complex, non-monotonic effect on the absolute confirmation bias, which was largest for low values (peaking at $\alpha = 0$), decreased, peaked again at intermediate values, and then steadily fell. In contrast, the normalized confirmation bias showed a simpler monotonic decrease as α increased (Fig. S6a-b, middle-left panels). The choice-trace accumulation rate (τ) also exhibited a non-monotonic effect, where its intermediate values produced the largest artifact (Fig. S6a-c, middle-right panels).Interestingly, the nature of the artifact from τ depended on its magnitude: low values primarily led to an underestimation of α_d, while high values primarily led to an overestimation of α_c. Finally, the artifactual bias increased monotonically as a function of the perseveration weight parameter (ϕ). This was driven by a corresponding monotonic increase in both the overestimation of α_c and the underestimation of α_d , with the relationship being asymmetric: perseveration ($\phi > 0$) generated a stronger effect than anti-perseveration ($\phi < 0$) (Fig. S6a-c, right panels).

Crucially, the asymmetry in the effect of ϕ implies that even a sample of participants with anti-perseverative tendencies (i.e., mean $\phi < 0$) can still produce a net spurious confirmation bias at the group level. To test this explicitly, we generated 2,000 simulations of the PSL model of experiment P2 based on the average MAP empirical parameters from P2, but with the ϕ parameter drawn from a uniform distribution between -9.25 and 8.75 (mean ϕ = -0.25, p=0.002; Fig. S7a). When this dataset was fitted with the hybrid model, we detected a significantly positive confirmation bias, for both the absolute (mean=0.03, p<0.001; Fig. S7b) and normalized metrics (mean=0.04, p<0.001; Fig. S7c).

Although our simulation was based on a specific set of parameters, they reveal that the generation of artifactual confirmation bias is not a monolithic effect. Instead, its magnitude and nature depend on a complex and often non-linear interplay between all parameters of the generative process. This highlights that subtle differences in behavior—such as the degree of choice stochasticity, the speed of value learning, or the precise rate at which choice history accumulates—can dramatically alter the likelihood that perseverative behavior will be misidentified as a confirmation bias.

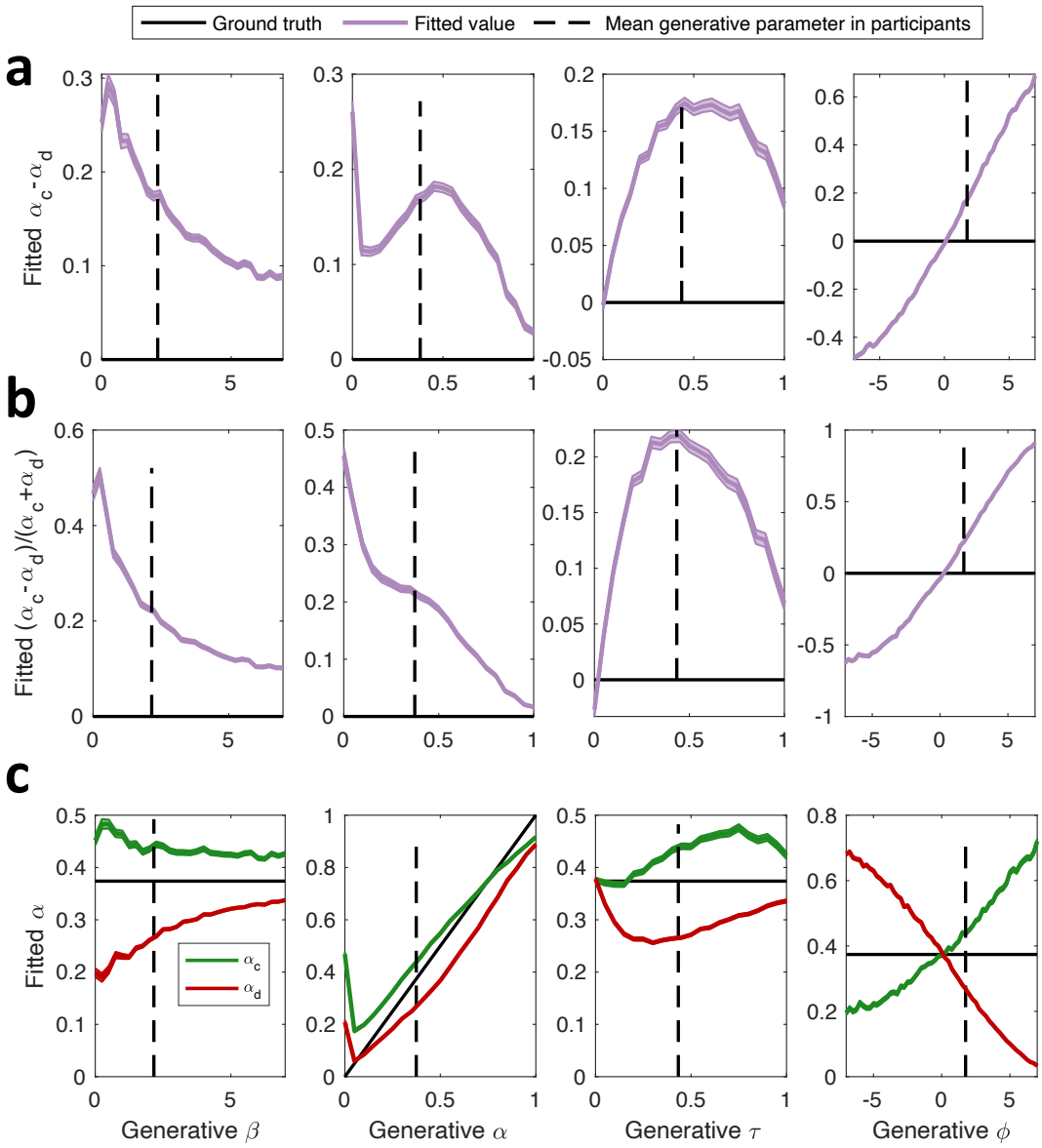

Figure S6: Parameter sweep showing how generative parameters from the perseveration model influence spurious confirmation bias. *The figure shows the results of a parameter sweep designed to isolate the influence of each generative parameter (β, α, τ, and ϕ). For each sweep one parameter was varied across a range of values, while the other three were held fixed at the mean values estimated from the PSL model fits of experiment P2.* ***(a)*** *Recovered absolute confirmation bias ($\alpha_c - \alpha_d$) as a function of the four generative parameters.* ***(b)*** *Same as (a), but for the normalized confirmation bias metric, $(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)$.* ***(c)*** *Disaggregation of the effects, showing the separately recovered confirmatory (α_c; in green) and disconfirmatory (α_d; in red) learning rates as a function of the generative parameters. In all panels, colored lines and shaded areas represent the mean ± SEM across 1001 simulations. Solid black horizontal lines represent the true generative values (0 for confirmation bias in a/b; the generative learning rate in c). Dashed black vertical lines represent the mean parameter value observed in the empirical data of experiment P2.*

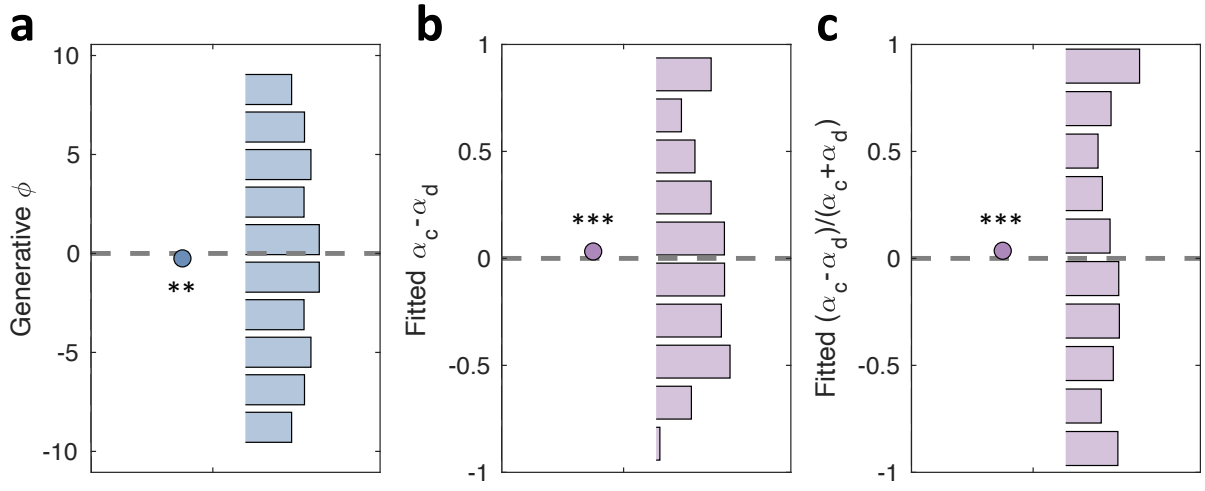

Figure S7: A population with a group-level negative perseveration can still produce a spurious confirmation bias. (a)*The distribution of the generative perseveration parameter (ϕ) from the PSL simulations, which follows a uniform distribution between -9.25 and 8.75, with a mean value of -0.25,.* ***(b)*** *The absolute confirmation bias detected when fitting these PSL simulations. The fitted bias is significantly positive (p<0.001).* ***(c)*** *Same as in (b), but for the normalized confirmation bias, which is also significantly positive (p<0.001).* In all panels, dots represent the mean parameter, and histograms represent the parameter distribution.

1.7. Bootstrapping procedure inflates Type I errors when performed with MAP estimation

To validate our bootstrapping procedure, we applied it to simulated datasets where the null hypothesis (H0) of symmetric learning was known to be true. We found that when using Maximum Likelihood Estimation (MLE), the procedure correctly failed to reject H0 (see SI 2.4). However, when using the biased Maximum a Posteriori (MAP) estimation, the procedure incorrectly rejected the null hypothesis for several individual experiments and at the meta-analytic level (see Fig S8). These results illustrate that using MAP estimation within our bootstrapping framework is susceptible to making a Type I error.

The reason for this inflated Type I error is a cascading bias introduced by the use of parameter priors in the MAP fitting procedure. The process begins when the PSL model is first fitted to the data to estimate the perseveration parameters under H0; as we established in the main text, the MAP procedure systematically underestimates the true ϕ parameter at this initial stage. The null distribution is then generated by simulating PSL agents using these *underestimated* ϕ parameters. When these new simulations are fitted with the hybrid model, the resulting spurious confirmation bias is smaller than the artifact that would have been produced by the *true*, higher level of perseveration, creating a null distribution that is biased towards zero. Finally, when the spurious confirmation bias from the original PSL data is compared

against this underestimated null distribution, it appears artificially large, leading to a spurious rejection of the null hypothesis.

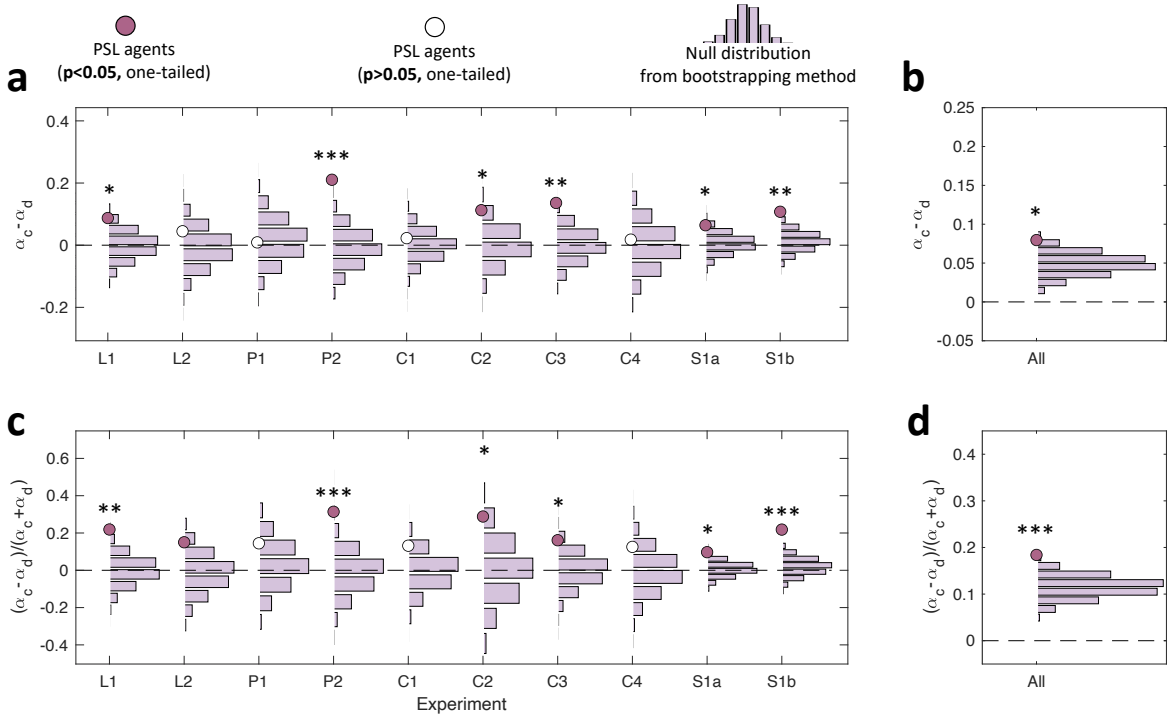

Figure S8: Testing confirmation bias in PSL agents against a null distribution of artefactual bias based on the MAP fitting procedure. ***(a)*** *Absolute confirmation bias for PSL simulations for each of the 10 experiments. Each violet histogram represents the null distribution of the mean artifactual bias, generated via the bootstrapping procedure using MAP estimation. The overlying dots represent the spurious confirmation bias obtained by fitting the original PSL agent simulations with the hybrid model. Stars indicate a significant (and incorrect) rejection of the null hypothesis (* p < 0.05; ** p < 0.01; *** p < 0.001).* ***(b)*** *Meta-analytic equivalent of panel (a), showing the mean absolute confirmation bias aggregated across all 10 experiments.* ***(c)*** *Same per-experiment analysis as panel (a), but for the normalized confirmation bias metric.* ***(d)*** *Meta-analytic result for the normalized confirmation bias, corresponding to panel (c).*

2. Supplementary MLE results

2.1. MLE Does Not Introduce Spurious Asymmetries in the Absence of Perseveration

To confirm that the Maximum Likelihood Estimation (MLE) procedure is less susceptible to the estimation biases observed with MAP, we repeated the control analysis from SI 1.3 using MLE at every step.

The procedure was identical: we first fitted the simple symmetric learning model (symmetric learning, no perseveration) to the empirical data using MLE to obtain a set of generative parameters. We then used these parameters to generate 1,001 simulated datasets per participant. Finally, we fitted these simulated datasets with the full Hybrid model using the MLE procedure and averaged the fitted parameters across their corresponding 1,001 simulations.

The results from this analysis contrast sharply with those from the MAP procedure. While we still found significant learning asymmetries in some individual experiments, there was no consistent direction: the bias was sometimes positive (absolute, $p<0.05$ in S1a and S1b; normalized, $p<0.05$ in S1a Fig. S9a-b) and sometimes negative (absolute, $p<0.05$ in P2,C1,C4; normalized, $p<0.05$ in P2,C1,C4,S1b Fig. S9a-b). Similarly, the recovered perseveration parameter (ϕ) was not consistently negative; it was significantly negative in some experiments (L1,L2,C1,S1a,S1b), non-significant in others (P1,C2,C3,C4), and even significantly positive in one case (P2).

Together, these findings show, while MLE is still prone to model mimicry issues, it is not prone to the same systematic estimation biases as the MAP procedure (see SI 1.3). It does not consistently generate a spurious confirmation bias from data that contains neither learning asymmetries nor perseveration, further validating its use in our main analyses.

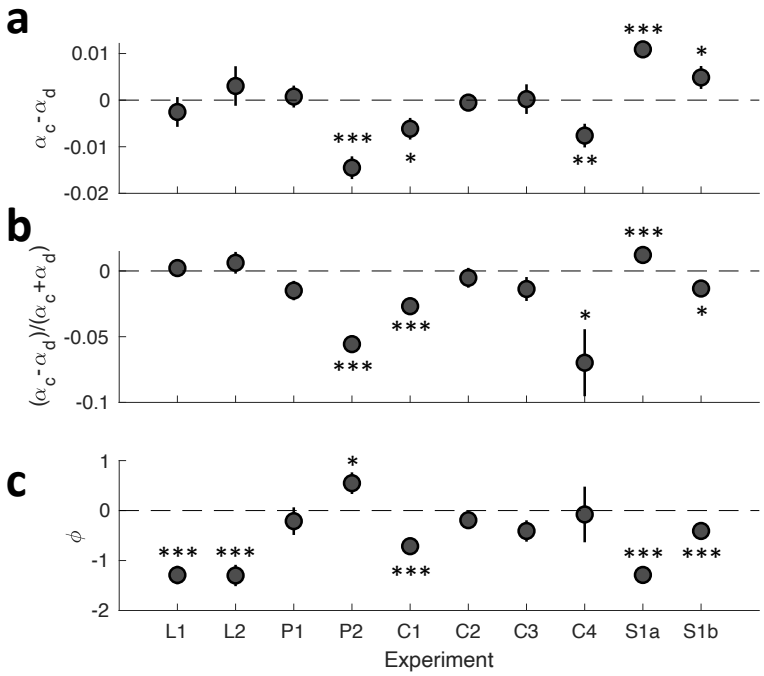

Figure S9: Reduced estimation bias in simulations using the MLE procedure. *The figure shows the results of fitting the Hybrid model with MLE to simulated data from a simple model with symmetric learning and no perseveration.* ***(a)*** *The fitted absolute confirmation bias is not consistently positive across experiments.* ***(b)*** *Same as (a), but for the normalized confirmation bias.* ***(c)*** *The fitted perseveration parameter (ϕ) is not consistently negative. Across all panels, the ideal unbiased estimate would be zero. Points and error bars represent the mean ± SEM across simulations. (* p < 0.05; ** p < 0.01; *** p < 0.001).*

2.2. Choice Perseveration Generates Spurious Confirmation Bias

When reproducing the results from Figure 2 using an MLE, we found similar results. The pure CB model detected a large confirmation bias for most experiments (absolute, all p's<0.05, except for P1; normalized, all p's<0.05), which was again significantly reduced when fitting the hybrid model (absolute, all p's<0.05 except L1,P2 and C4; normalized, all p's<0.05). However, when fitting the data with the hybrid model using MLE, only a minority of experiments produced a significant confirmation bias (absolute, p<0.05 for P2,C2 and C4; normalized, p<0.05 for L1,P2,C1 and C4), as opposed to a majority when using the MAP procedure. Similarly, simulations of PSL agents produced a significant confirmation bias in some experiments (absolute, p<0.05 in P1, P2, C2, S1a and S1b; normalized, p<0.05 in L1, P1, P2, C2, S1a and S1b), again, less experiments than when using the MAP procedure. Finally, this spurious confirmation bias was lower than the one obtained when fitting the data to the hybrid model in a minority of experiments (absolute, p<0.05 in P2, C2 and C4; normalized, p<0.05 in P2, C1 and C4).

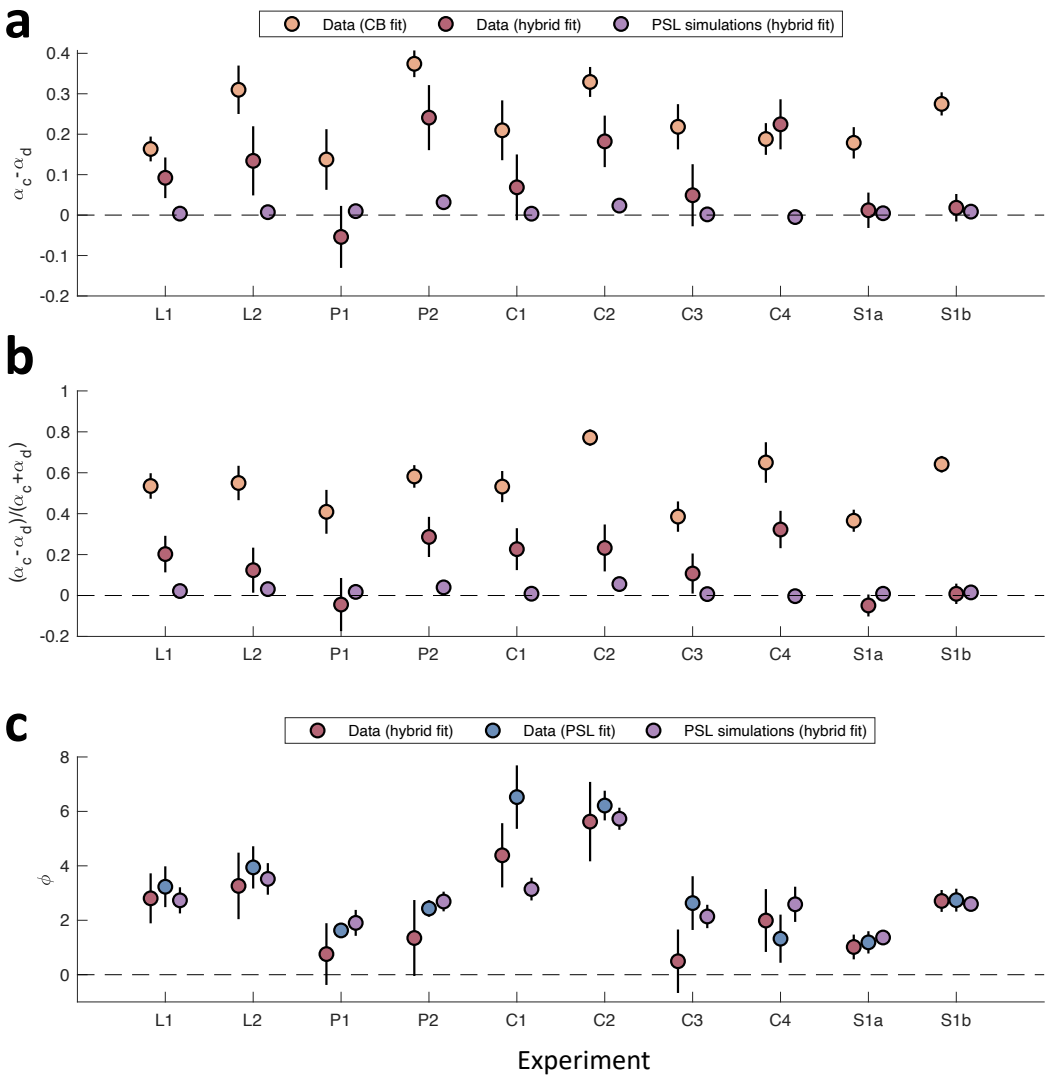

Figure S10: Confirmation bias and perseveration parameters estimated from empirical data and simulations fitted with MLE. (a) *Absolute confirmation bias ($\alpha_c - \alpha_d$) estimated from fitting participant data with the Pure CB model (**yellow points**) and the Hybrid model (**red points**). Additionally, **violet points** show the spurious confirmation bias recovered when fitting PSL simulations with the Hybrid model.* ***(b)*** *Same as (a), but for the normalized confirmation bias, defined as $(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)$.* ***(c)*** *The perseveration parameter (ϕ) estimated from participant data using the Hybrid model (**red points**) and the PSL model (**blue points**). **Violet points** show the ϕ values recovered from fitting the Hybrid model to simulations of PSL agents, whose original generative parameters are shown in blue. Points and error bars represent mean ± SEM across participants/simulations.*

2.3. Asymptotic Dissociation of Perseveration and Confirmation Bias in Large Datasets

Continuing the asymptotic analysis from the main text (Figure 4), we examined the estimation of the perseveration parameters, ϕ and τ when using the MLE procedure. The perseveration parameter (ϕ) was initially underestimated for simulations with very few trials (<200 trials) but converged quickly to its true generative value for datasets with more trials (Fig. S11a). In contrast, the choice-trace accumulation rate (τ) showed a more complex, non-monotonic estimation pattern. It was overestimated in datasets with a few trials, converging to its true generative value in datasets of around 300-400 trials. However, it was then underestimated for a subsequent range of larger datasets, only achieving final convergence for datasets larger than 9,000 trials (Fig. S11b).

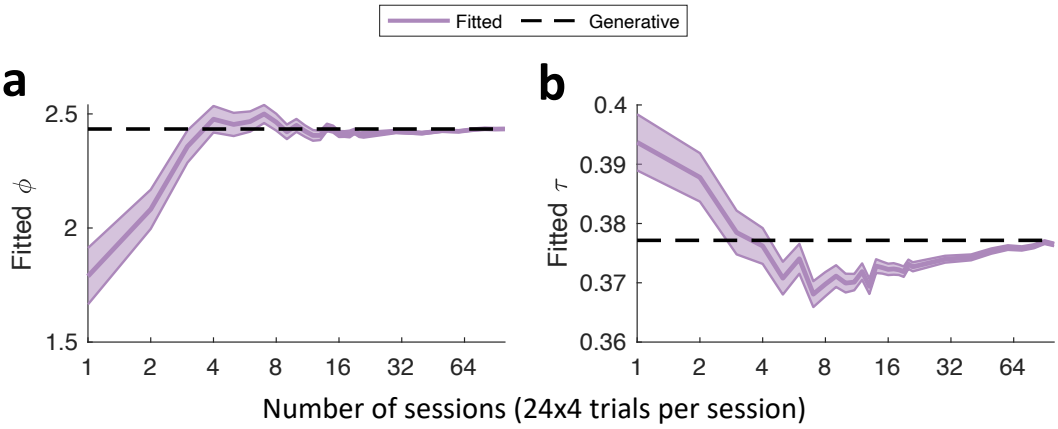

Figure S11: Misestimation of perseveration parameters fitted with MLE is reduced with dataset size. *The figure shows parameter estimates obtained by fitting the Hybrid model to simulations of PSL agents, as a function of the number of sessions in the simulated dataset (where each session encompasses 4 blocks of 24 trials). The simulations were based on the task in the second experiment (full feedback condition) from Palminteri et al. (2017* ***(a)*** *Fitted perseveration parameter (ϕ). The parameter is underestimated in small datasets (<200 trials) but quickly converges to its true generative value.* ***(b)*** *Fitted choice-trace accumulation rate (τ). The parameter shows a non-monotonic pattern: it is overestimated in small datasets, converges around 300-400 trials, is then underestimated, and only achieves final convergence for very large datasets (>9,000 trials). In all panels, solid lines and shaded areas represent the mean ± SEM across 200 simulations. Dashed lines represent the true generative parameter value used in the simulations. Gray boxes highlight dataset sizes for which the recovered parameter is significantly different from the generative value.*

2.4. Bootstrapping procedure correctly rejects null distribution for PSL simulations

To validate our bootstrapping procedure and confirm that it correctly controls for Type I errors, we applied it to simulated datasets where the null hypothesis (H0) of symmetric learning was known to be true. Specifically, we used the same PSL model simulations from the main text (Figure 2) as our "empirical" data and applied the full bootstrapping analysis (fitting the data with our PSL models, generating PSL simulations, fitting them with the hybrid model and bootstrapping the results to generate the null hypothesis distributions). We found that when using Maximum Likelihood Estimation (MLE), the procedure correctly failed to reject H0 for all individual experiments (all $p > 0.05$; Fig. S12a,c) and also at the meta-analytic level (absolute, $p=0.73$; normalized, $p=0.93$; Fig. S12b,d). The effectiveness of the procedure is well-illustrated by the C2 experiment simulations. While a direct fit of the hybrid model to these PSL simulations produced a spurious but statistically significant confirmation bias, our bootstrapping procedure correctly showed that this bias was not significantly greater than what would be expected under the null distribution, and therefore correctly failed to reject H0. This highlights how the procedure can effectively distinguish between a statistically significant but spurious effect and a genuine one.

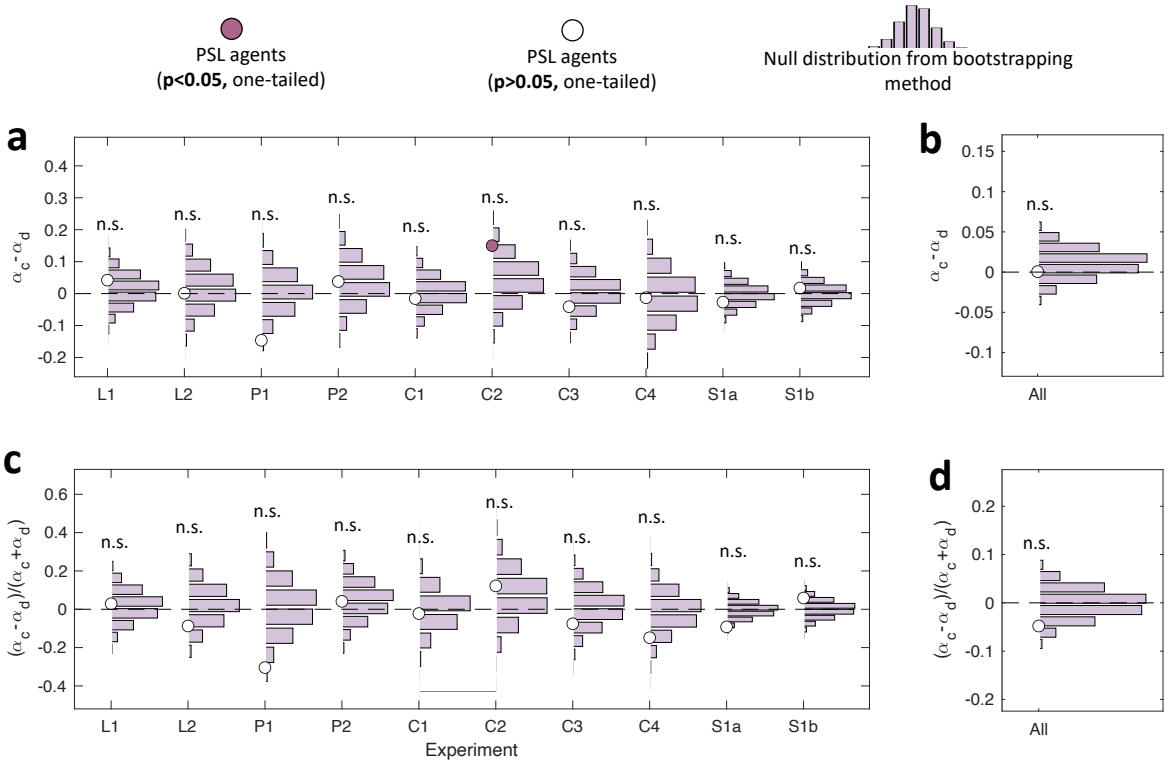

Figure S12: Testing confirmation bias in PSL agents against a null distribution of artefactual bias based on the MLE fitting procedure. (a) *Absolute confirmation bias for PSL simulations for each of the 10 experiments. Each violet histogram represents the null distribution of the mean artifactual bias, generated via the bootstrapping procedure using MLE. The overlying dots represent the spurious confirmation bias obtained by fitting the original PSL agent simulations with the hybrid model. Stars indicate a significant (and incorrect) rejection of the null hypothesis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).* ***(b)*** *Meta-analytic equivalent of panel (a), showing the mean absolute confirmation bias aggregated across*

all 10 experiments. ***(c)*** *Same per-experiment analysis as panel (a), but for the normalized confirmation bias metric*. ***(d)*** *Meta-analytic result for the normalized confirmation bias, corresponding to panel (c)*.

2.5. Contribution of Generative Parameters to Spurious Confirmation Bias under MLE

To better understand the mechanics of the artifact when using MLE, we conducted a parameter sweep analysis following the same procedure as in SI 1.7. This allowed us to characterize which parameters of the PSL model most strongly drive the emergence of the artifactual confirmation bias when using fitting procedure that is less biased by parameter priors.

This analysis again revealed complex relationships between the generative parameters and the artifactual bias. The relationship with the generative parameters was broadly the same as that observed when using MAP estimation, but with a few key differences. First, the large bias for low values of the choice inverse temperature (β) was driven primarily by an underestimation of the disconfirmatory learning rate (α_d), but not by an overestimation of the confirmatory learning rate (α_c) (Fig. S13a-c, left panels). Second, extreme values of the choice-trace accumulation rate (τ) produced a *disconfirmatory* bias, while intermediate values still produced a confirmation bias (Fig. S13a-c, middle-right panels). Finally, the artifactual bias as a function of the perseveration weight parameter (ϕ) was more asymmetric than the one observed when using MAP: perseveration ($\phi > 0$) generated a strong effect, while anti-perseveration ($\phi < 0$) generated a very small effect and plateaued quickly (Fig. S13a-c, right panels).

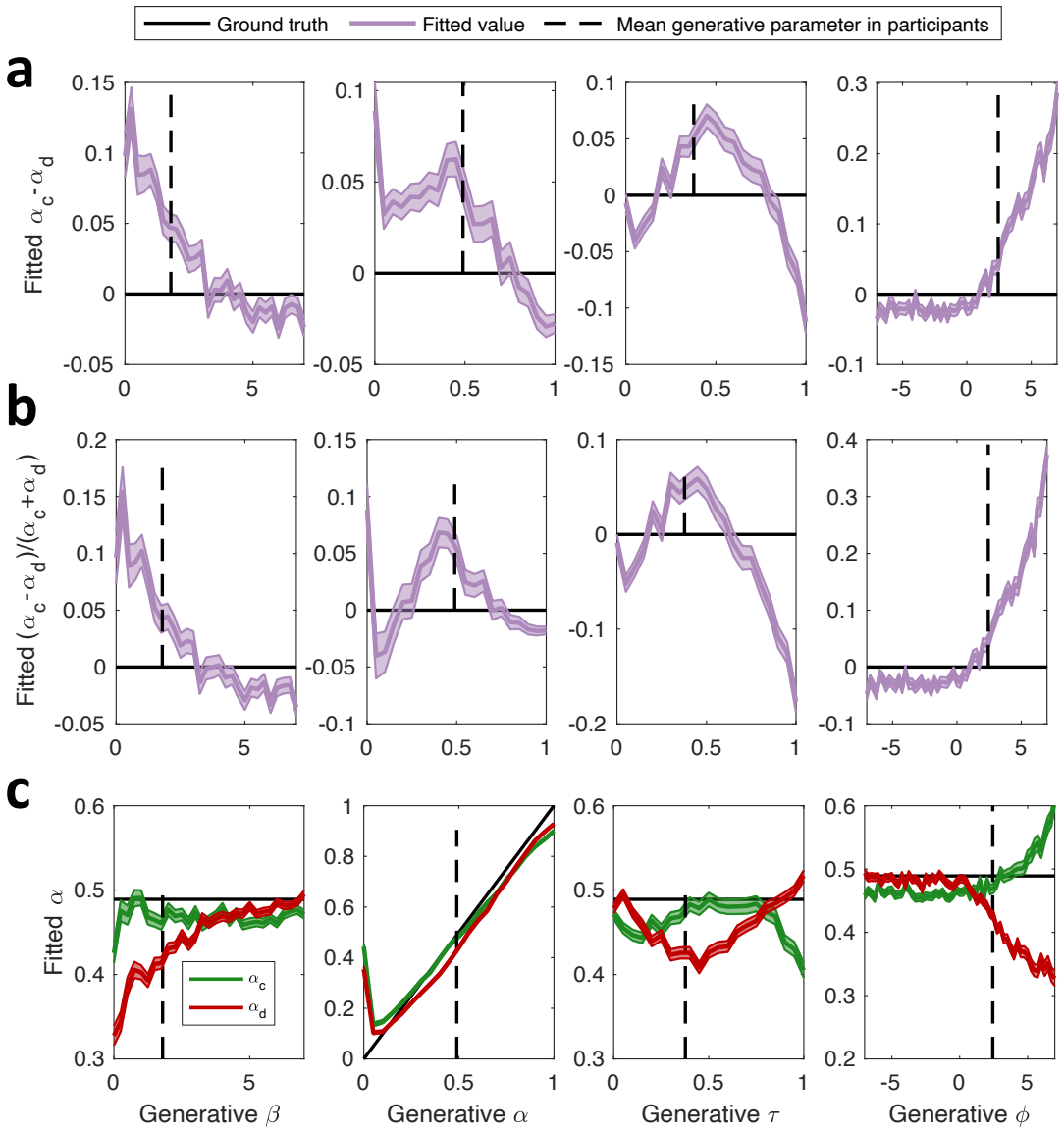

Figure S13: Parameter sweep showing how generative parameters from the perseveration model influence spurious confirmation bias. *The figure shows the results of a parameter sweep designed to isolate the influence of each generative parameter (β, α, τ, and ϕ). For each sweep one parameter was varied across a range of values, while the other three were held fixed at the mean values estimated from the PSL model fits of experiment P2.* ***(a)*** *Recovered absolute confirmation bias ($\alpha_c - \alpha_d$) as a function of the four generative parameters.* ***(b)*** *Same as (a), but for the normalized confirmation bias metric, $(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)$.* ***(c)*** *Disaggregation of the effects, showing the separately recovered confirmatory (α_c; in green) and disconfirmatory (α_d; in red) learning rates as a function of the generative parameters. In all panels, colored lines and shaded areas represent the mean ± SEM across 1001 simulations. Solid black horizontal lines represent the true generative values (0 for confirmation bias in a/b; the generative learning rate in c). Dashed black vertical lines represent the mean parameter value observed in the empirical data of experiment P2.*

3. Supplementary behavioral signature results

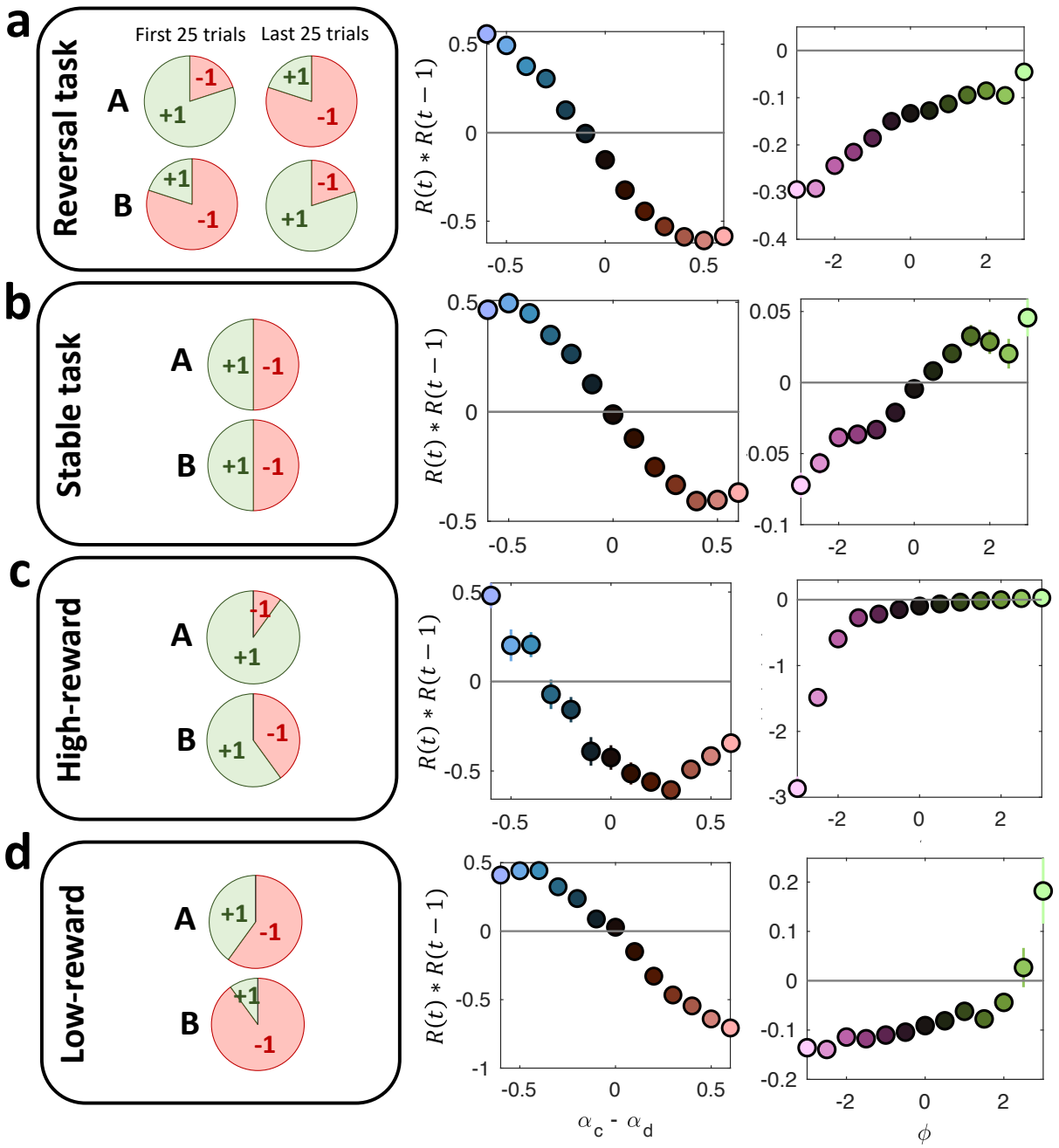

Figure S14: Robustness of the Katahira (2018) regression signature across different task designs. *Test for the robustness of the regression-based signature proposed by Katahira (2018) across four additional typical bandit task designs. The analysis is performed in the same way as the one in the main text. All simulations consist of a single 50-trial block.* ***(a)*** *Reversal Task. We simulated a task with 80%/20% reward probabilities, reversed at trial 26. All tested perseveration levels produced a negative interaction that mimics the signature of a confirmation bias.* ***(b)*** *Stable Task. We simulated a task with two 50%/50% bandits without reversals. Negative perseveration produced a negative interaction that mimics the signature of a confirmation bias.* ***(c)*** *High-Reward Stable Task. We simulated a task with 90%/60%*

reward probabilities without reversals. Negative perseveration produced a negative interaction that mimics the signature of a confirmation bias. ***(d)*** *Low-Reward Stable Task. We simulated a task with 40%/10% reward probabilities without reversals. Perseveration levels below 2.5 produced a negative interaction that mimics the signature of a confirmation bias.*