

SUPPLEMENTARY INFORMATION

Distinguishing learning asymmetry from perseveration in behavioral models

Juan Vidal-Perez^{1,2*}, Raymond J. Dolan^{1,2}, Rani Moran^{1,3*}

***Corresponding authors.** Email: rani.moran@gmail.com, juan.perez.21@ucl.ac.uk

Affiliations:

¹ Max Planck Centre for Computational Psychiatry and Ageing, University College London, Russell Square House, WC1B 5EH, England

² Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3BG, England

³ Department of Psychology, School of Biological and Behavioural Sciences, Queen Mary University of London, London, United Kingdom

Table of Contents

SUPPLEMENTARY NOTES	3
Supplementary note 1: MAP results	3
1.1. Choice perseverance masquerades as spurious normalized confirmation bias	3
1.2. The Hybrid model underestimates perseverance in PSL agents.....	3
1.3. Choice Perseveration Generates Spurious Confirmation Bias in Expanded Models	4
1.4. Parameter trade-offs in hybrid model fits and choice repetition predictions.....	7
1.5. MAP Estimation Introduces Spurious Asymmetries in the Absence of Perseveration	8
1.6. Confirmation bias produces little spurious positive perseverance	9
1.7. MAP estimation generates spurious confirmation bias when using a wider prior for the perseveration parameter ($N(0,5)$)	11
1.8. Confirmation bias overestimation varies as a function of the generative confirmation bias.....	13
1.9. Asymptotic Dissociation of Perseveration and Confirmation Bias in Large Datasets	14
1.10. Contribution of different parameters to the spurious confirmation bias	17
Supplementary note 2: MLE results.....	20
2.1 MLE Does Not Introduce consistent spurious Asymmetries in the Absence of Perseveration	20
2.2. Choice Perseveration Generates Spurious Confirmation Bias	21
2.3 Confirmation bias overestimation varies as a function of the generative confirmation bias.....	23
2.4 Confirmation bias produces little spurious positive perseverance	24
2.5 Asymptotic Dissociation of Perseveration and Confirmation Bias in Large Datasets	26
2.6 Group-level test against a null distribution of normalized spurious confirmation bias	29
2.7 Contribution of Generative Parameters to Spurious Confirmation Bias under MLE	30
2.8 A group with average negative perseverance can produce a significant spurious confirmation bias	32
Supplementary note 3: New hypothesis-testing procedures	32
3.1 Individual-level test against a null distribution of spurious confirmation bias	33
3.2 Group-level test against a null distribution of spurious confirmation bias (APPS Test)	36
Supplementary note 4: Behavioral signature results.....	42
SUPPLEMENTARY TABLES	46
A Hybrid Model Reduces, but Does Not Eliminate, Measured Confirmation Bias in Empirical Data	46
Choice perseverance masquerades as spurious confirmation bias.....	47
SUPPLEMENTARY METHODS	49
The Modeler's Handbook: A Guide to Addressing Mimicry Problems	49
Method I. Is there a model mimicry/fitting bias problem?	50
Method II. Type I Error Validation. Does a test lead to an inflated false positive rate?	51
Method III. The Anti-Participant Permutation Simulations test (APPS test).....	52

SUPPLEMENTARY NOTES

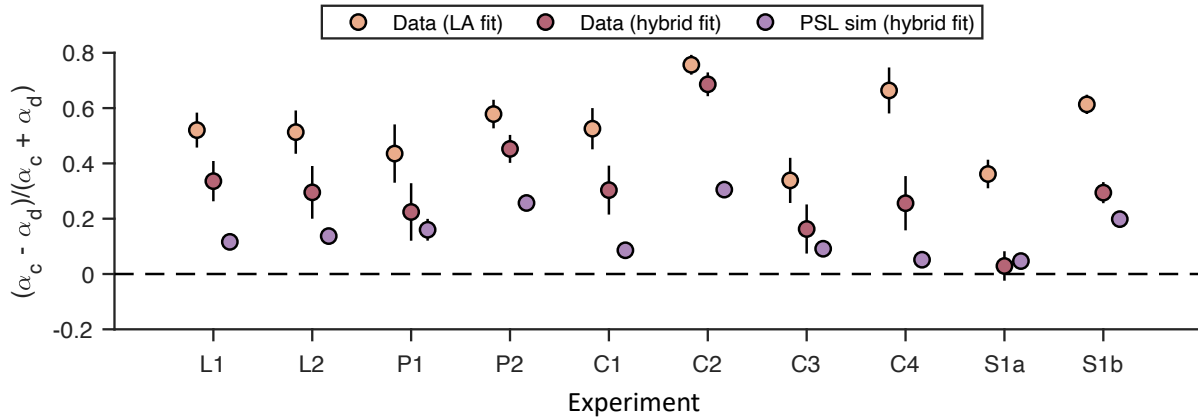
Supplementary note 1: MAP results

1.1. Choice perseveration masquerades as spurious normalized confirmation bias

In addition to the absolute difference in learning rates ($\alpha_c - \alpha_d$), we evaluated the susceptibility of the normalized confirmation bias metric (defined as $(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)$) to perseverative confounds. This is another commonly used metric for value learning asymmetries, which takes into account the variability in overall learning rate across participants (9,30).

We first examined the normalized confirmation bias in empirical data by comparing a Pure LA model (which ignores perseveration) to the Hybrid model across the same 10 experiments as in Fig 2 in the main text. For the Pure LA model, a significant normalized confirmation bias was detected across all experiments (all $p < 0.05$; Supplementary Fig. 1 yellow points). When fitting the Hybrid model, this bias remained significant in most experiments (all $p < 0.05$ except for C3 and S1a; Supplementary Fig. 1 red points), but its magnitude was significantly reduced compared to the Pure LA model (significant decrease, $p < 0.05$ for all experiments except for C3).

Our simulation and re-fitting analysis revealed that the normalized metric is also vulnerable to model-mimicry artifacts. When fitting the hybrid model to behavior generated by perseverative symmetrically learning (PSL) agents where the ground truth bias is zero, the spurious normalized confirmation bias was statistically significant across all ten analyzed datasets ($p < 0.05$; Supplementary Fig. 1 violet points).



Supplementary Figure 1: Normalized confirmation bias $(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)$ estimated from fitting participant data with the Pure LA model (yellow points) and the Hybrid model (red points). Additionally, violet points show the spurious normalized confirmation bias recovered when fitting PSL simulations with the Hybrid model.

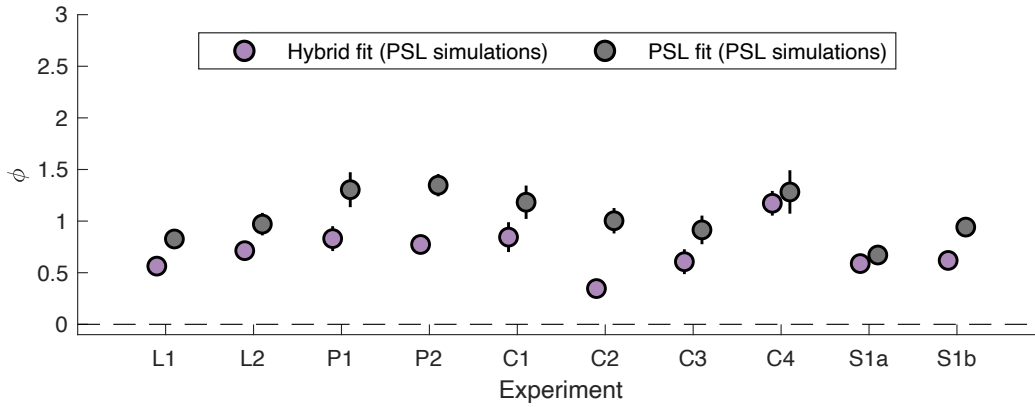
1.2. The Hybrid model underestimates perseveration in PSL agents

To provide further evidence for the tradeoff between perseveration and learning asymmetries reported in the main text (Figure 2b), 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 Supplementary Figure 2, 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.



Supplementary Figure 2: Comparison of perseveration parameter (ϕ) when fitting PSL model simulations with either PSL model or the Hybrid model. The Hybrid model (violet) consistently yields a significantly lower estimate of ϕ than the PSL model (gray). Points and error bars represent the mean $\pm$ SEM across simulations for each experiment.

1.3. 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:

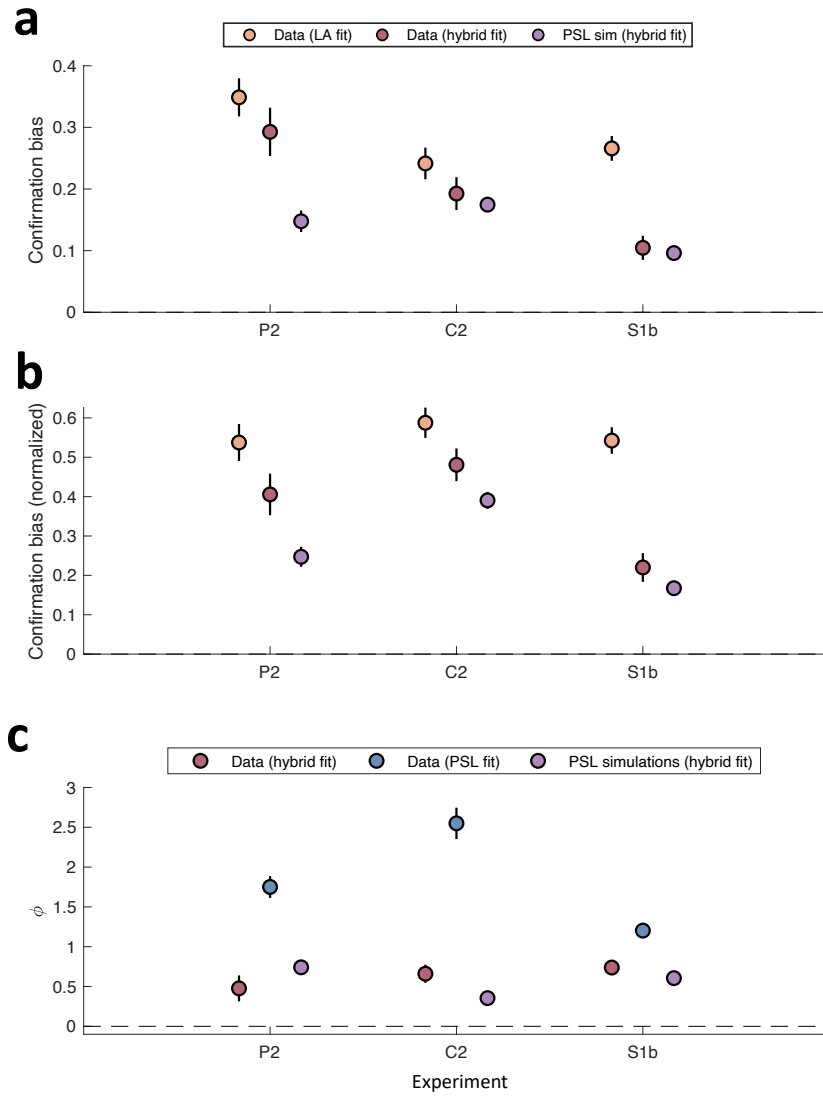
$$Q_{t+1}(C) = \begin{cases} Q_t(C) + \alpha_c^+ \cdot \delta_t(C) & \text{if } \delta_t(C) > 0 \\ Q_t(C) + \alpha_c^- \cdot \delta_t(C) & \text{if } \delta_t(C) < 0 \end{cases}$$

And unchosen options are updated as:

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

The Expanded LA 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 LA 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.

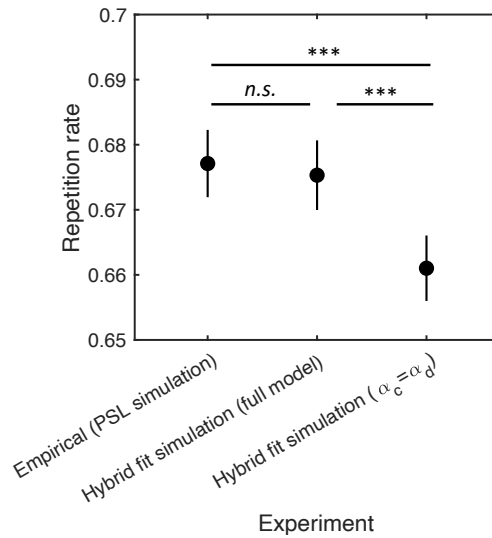


Supplementary Figure 3: Confirmation bias and perseveration parameters estimated with Expanded Models. **(a)** Absolute confirmation bias $[(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/2]$ estimated from fitting participant data with the Expanded LA model (yellow points) and the Expanded Hybrid model (red points). Additionally, violet points show the spurious confirmation bias recovered when fitting PSL simulations with the Expanded Hybrid model. **(b)** $(\alpha_c^+ + \alpha_u^- - \alpha_c^- - \alpha_u^+)/(\alpha_c^+ + \alpha_u^- + \alpha_c^- + \alpha_u^+)$. **(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 $\pm$ SEM across participants/simulations for the three full-feedback experiments.

1.4. Parameter trade-offs in hybrid model fits and choice repetition predictions

Figure 2 in the main text highlighted a spurious confirmation bias and its tradeoff with perseveration. We hypothesized that, when fitting the hybrid model to Perseverative Symmetric-Learning (PSL) simulations, the model's perseveration weight (ϕ) is underestimated (often due to shrinkage from the parameter prior), so a spurious confirmation bias ($\alpha_c > \alpha_d$) emerges to avoid under-predicting the empirical rate of choice repetition.

We tested this intuition using a three-step simulation and ablation analysis. First, we created PSL agents based on the (MAP fitted) PSL model parameters fitted to the actual participant data, and we simulated synthetic behavior for these agents (This step is identical to what we did for Fig. 2) We calculated the average choice repetition rate for these simulations (pooling across all experiments), which served as our "empirical" baseline for this analysis. Next, we fitted these purely perseverative PSL simulations with the full hybrid model. Hybrid model simulations based on the recovered hybrid parameters (recall from main text Fig. 2a that these parameters featured a spurious confirmation bias) successfully captured the empirical choice repetition rates (Supplementary Fig. 4, left vs middle points; $p=0.33$). Finally, we performed an ablation to determine if the spurious bias helped predicting the "empirical" repetition rate. Towards this aim, we took the recovered hybrid parameters from the previous step but forced the learning rates to be symmetric by setting both to their mean ($new\ \alpha_c \ \& \ \alpha_d = (\alpha_d + \alpha_c)/2$). Hybrid simulations, based on the ablated parameters, resulted in a significant drop in choice repetition compared to both the empirical baseline (Supplementary Fig. 4 right vs left points) and the full hybrid model (Supplementary Fig. 4 middle vs right points) simulations (both p 's < 0.001). These results provide direct evidence for a compensatory trade-off. An underestimation of ϕ , is accompanied by a spurious confirmation bias to bring the choice repetition back toward empirical levels. Because both positive confirmation bias and choice perseveration predict an increase in choice repetition, the fitting procedure incorrectly partitions the variance from a single habitual mechanism into two separate, misestimated parameters.



Supplementary Figure 4: Spurious confirmation bias helps the hybrid model to account for repetition rates. This figure illustrates the average choice repetition rate across three simulation conditions to highlight the compensatory relationship between perseveration and confirmation bias. The left point represents the average choice rate of PSL simulations based on participants' parameters (empirical baseline). The middle point shows the repetition rate recovered by fitting these PSL empirical simulations with the hybrid model and using these hybrid parameters to simulate the hybrid model. The right point shows the repetition rate after ablating the recovered confirmation bias in these hybrid fits (by forcing the learning rates to be symmetric). The significant drop in choice repetition following ablation ($p < 0.001$) demonstrates that the spurious bias is partly responsible accurate choice repetition predictions of hybrid model fits. Points represent average choice repetition rates (of all experiments pooled together) and error bars represent the s.e.m. (n.s., $p > 0.05$; *** $p < 0.001$)

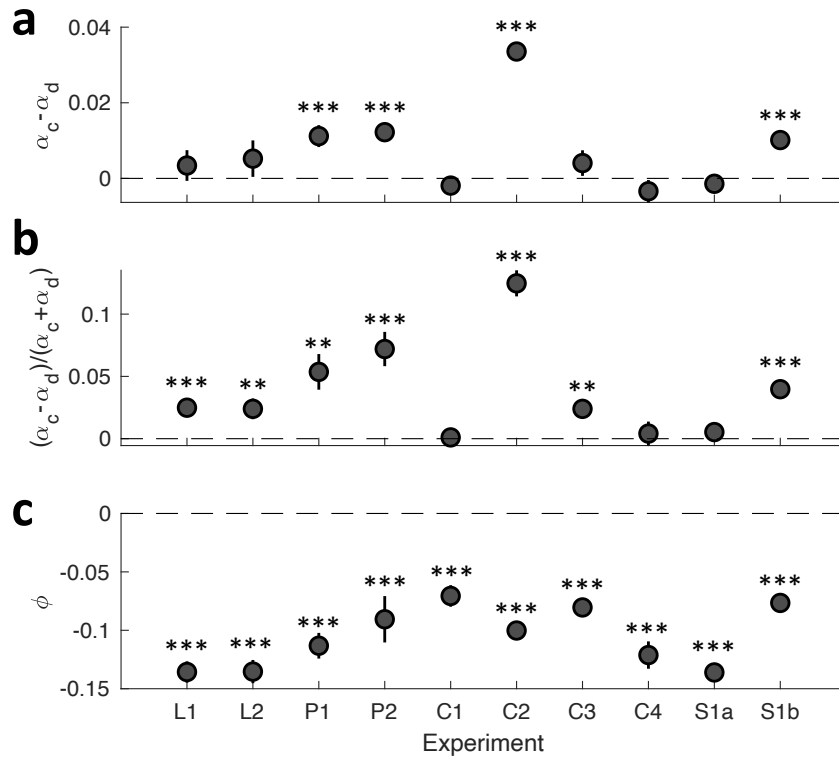
1.5. MAP Estimation Introduces Spurious Asymmetries in the Absence of Perseveration

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 LA 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; Supplementary Fig. 4a-b). Furthermore, the perseveration parameter (ϕ) was consistently estimated to be significantly *below* zero across all experiments (all p 's < 0.001 ; Supplementary Fig. 4c).

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.



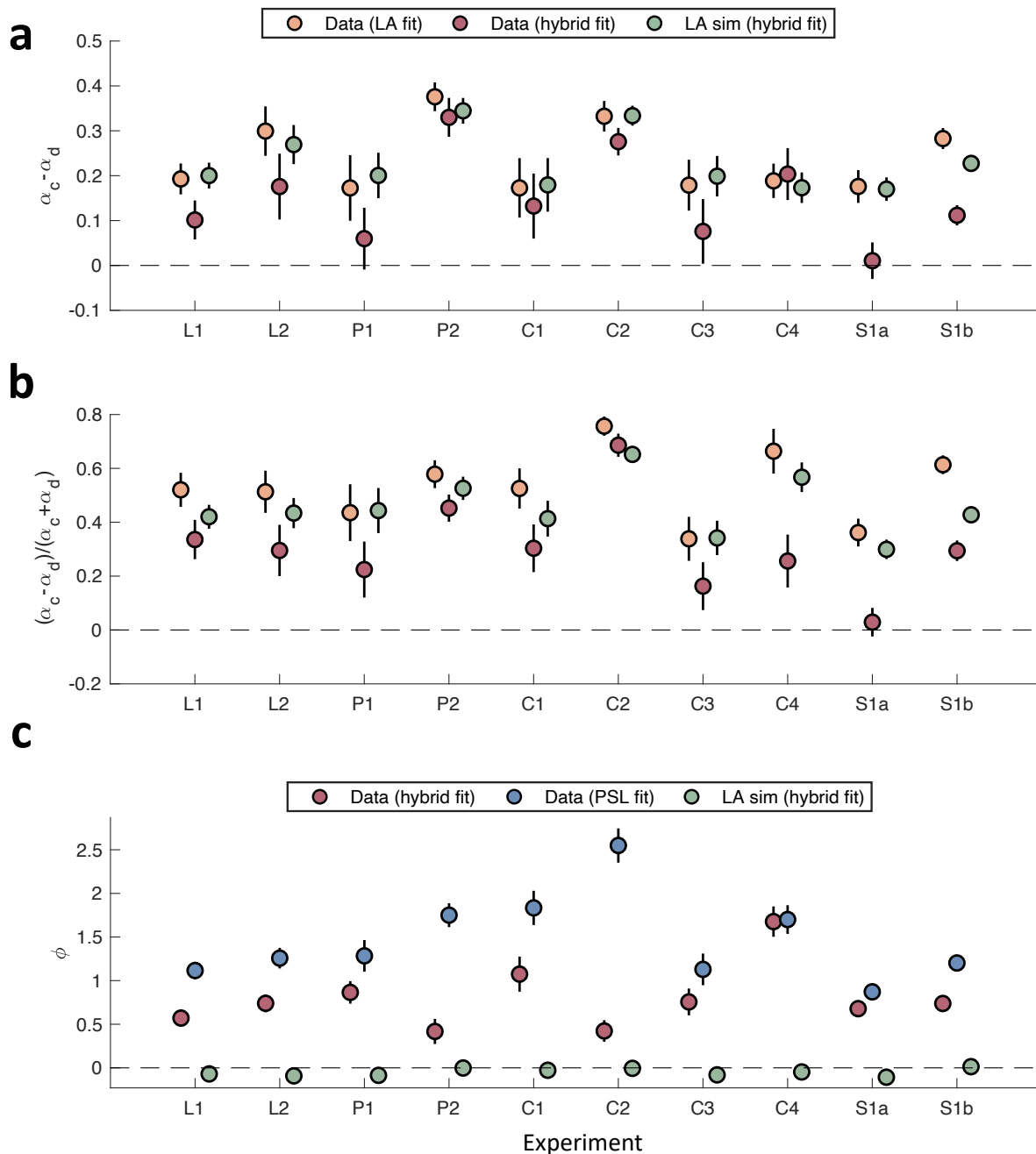
Supplementary Figure 5: Spurious parameters recovered from simulations of a simple symmetric learning model. The figure shows the results of fitting the Hybrid model to simulated data from a simple model with symmetric learning and no perseveration. **(a)** Absolute confirmation bias fitted from the 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 $\pm$ SEM across simulations. (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

1.6. Confirmation bias produces little 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 learning asymmetries could lead to the spurious detection of perseveration. To this end, we simulated 1001 datasets per participant from our LA model (using the MAP-estimated empirical parameters) and subsequently fitted these datasets with our hybrid model.

In nearly all cases, the fitted confirmation bias parameters did not differ significantly from the generative empirical parameters, when measured in absolute terms (yellow vs green points in Supplementary Fig. 6a; all p 's > 0.05 , except for P2 and S1b $p < 0.001$), but were significantly lower in most experiments when measured in normalized terms (yellow vs green points in Supplementary Fig. 6b; normalized, $p < 0.05$ for L1, L2, P2, C1, C2, S1a and S1b). We did not find a consistently positive spurious 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) (Supplementary Fig. 6c). Together, these results suggest that the confound is largely

unidirectional. Whereas perseveration can be misattributed as a confirmation bias, a true confirmation bias is generally recovered accurately by the hybrid model and not misattributed as a positive perseveration.



Supplementary Figure 6: Confirmation bias and perseveration parameters estimated from empirical data and LA simulations fitted using MAP. (a) Absolute confirmation bias estimated from fitting participant data with the Pure LA model (yellow points) and the Hybrid model (red points). Additionally, green points show the confirmation bias recovered when fitting LA 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 LA agents, showing that no spurious positive perseveration parameter emerges for most experiments. Points and error bars represent mean $\pm$ SEM across participants/simulations.

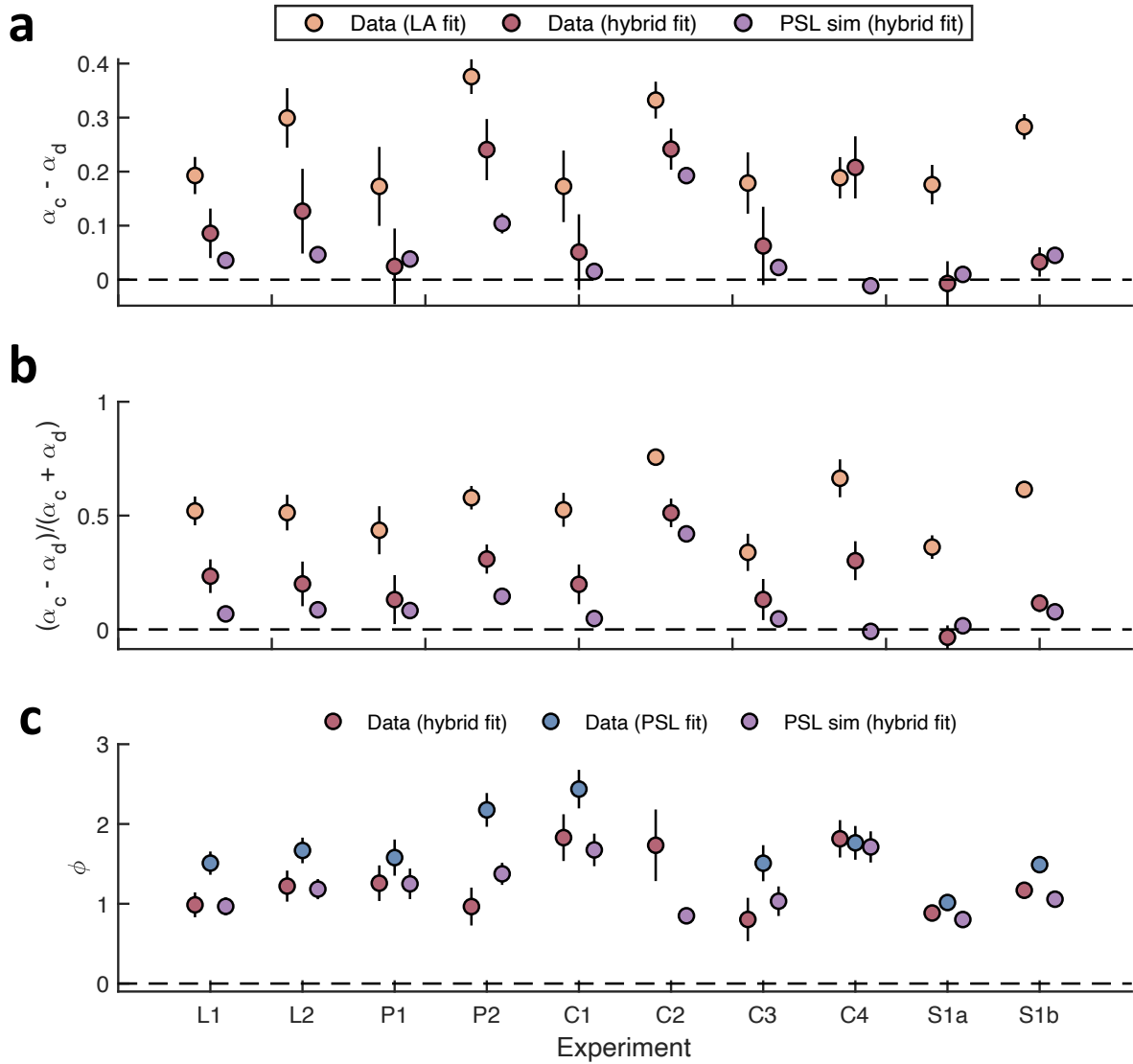
1.7. MAP estimation generates spurious confirmation bias when using a wider prior for the perseveration parameter ($N(0,5)$)

In the main text, we used a narrow normal distribution ($N(0,1)$) as a prior for the perseveration parameter (ϕ), consistent with experiments L1-2, P1-2 and C1-4 in our dataset. However, other researchers, most notably Sugawara and Katahira (experiment S1a-b), employed a wider prior with a variance of 5 ($N(0,5)$). It has been argued that a failure to detect a significant confirmation bias reported in Sugawara and Katahira's original study is a direct consequence of this wider prior (Palminteri, 2023). Thus, it is plausible that a wider prior mitigates shrinkage of the perseveration parameter and thus allows for a cleaner dissociation of the two mechanisms.

To test whether a wider prior eliminates the spurious learning asymmetries described in Figure 2, we repeated our full simulation and re-fitting analysis using a $N(0, 5)$ prior. Our results indicate that while a wider prior reduces the magnitude of observed effects, the core model mimicry problem persists.

First, we examined the behavior of PSL agents under this wider prior. We found that the hybrid model still recovered a significant spurious confirmation bias (Supplementary Fig. 7a-b violet points; absolute: $p < 0.05$ in all experiments but C4; normalized: $p < 0.05$ in all experiments, except for C4). Using the wider prior, we compared confirmation bias estimates obtained from the Pure learning asymmetries (Pure LA) model and the Hybrid model. Similar to our main analysis, inclusion of the perseveration parameter reduced the estimated bias (absolute: $p < 0.05$ in all experiments except for C1, C2, and C4; normalized: $p < 0.05$ in all experiments). Consistent with the underestimation observed in the main text, there was a significant decrease in the estimated value when moving from the PSL model to the Hybrid model fits of empirical data ($p < 0.05$ in all experiments except for C4; Supplementary Fig. 7c blue vs red points). Furthermore, in our simulation analysis, the recovered parameters were systematically lower than the true generative values ($p < 0.05$ in all experiments except for C4; blue vs violet points in Supplementary Fig. 7c).

Together, these findings reveal that the use of a wider $N(0,5)$ prior does not eliminate the tendency of the hybrid model to misattribute perseverative variance to learning rate asymmetries. Even under a $N(0, 5)$ prior, purely perseverative behavior can masquerade as confirmation bias.



Supplementary Figure 7: Confirmation bias and perseveration parameters estimated from empirical data and simulations using a fitting prior of $N(0,5)$ for the perseveration parameter. (a) Absolute confirmation bias ($\alpha_c - \alpha_d$) estimated from fitting participant data with the Pure LA 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. A spurious confirmation bias is detected across most experiments even when the generative model has no learning asymmetry. **(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. The recovered violet points are systematically lower than the generative blue points, indicating underestimation. Points and error bars represent mean $\pm$ SEM across participants/simulations.

Additionally, we compared conclusions, regarding the 10 datasets, derived from the original MAP hybrid model fitting (using both $N(0,1)$ and $N(0,5)$ priors) against those obtained with our proposed Anti-Participant Permutation Simulation (APPS) method (Supplementary Fig. 8; all conclusions are based on one-tailed t-tests for a positive confirmation bias). We found that the wider $N(0, 5)$ prior is more conservative than the narrow prior, particularly when evaluating the absolute confirmation bias, where it produced a change in conclusion in 3 out of 10 experiments (L2, C1 and S1b; Supplementary Fig. 8a, top vs middle rows). In normalized terms, it produced a change in conclusion in 2 out of 10 experiments (P1 and C3; Supplementary Fig. 8b, top vs middle rows). However, the wider prior does not simply act as a proxy for our statistical control. Experiment L1 showed a significant absolute confirmation bias in the hybrid model when using the wider prior (Supplementary Fig. 8a, middle vs bottom rows), but did not reach significance when tested against our empirically-grounded null distribution. Conversely, 3 experiments that showed a significant normalized confirmation bias under the wider prior failed to reach significance when using our APPS method (L2, C1, and S1b; Supplementary Fig. 8b, top vs middle rows).

a	Absolute confirmation bias $[\alpha_c - \alpha_d]$										
	L1	L2	P1	P2	C1	C2	C3	C4	S1a	S1b	All
Original (N(0,1))	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓
Original (N(0,5))	✓	✗	✗	✓	✗	✓	✗	✓	✗	✗	✓
New method conclusion	✗	✗	✗	✓	✗	✓	✗	✓	✗	✗	✓

b	Normalized confirmation bias $[(\alpha_c - \alpha_d)/(\alpha_c + \alpha_d)]$										
	L1	L2	P1	P2	C1	C2	C3	C4	S1a	S1b	All
Original (N(0,1))	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓
Original (N(0,5))	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	✓
New method conclusion	✓	✗	✗	✓	✗	✓	✗	✓	✗	✗	✓

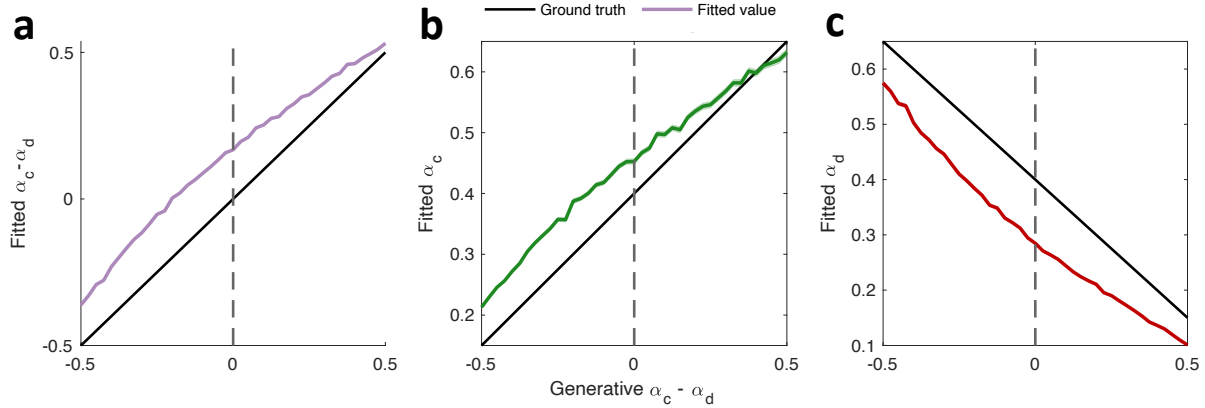
Supplementary Figure 8: Comparison between the conclusions drawn from the original method (Hybrid model with MAP) using a narrow prior ($N(0,1)$, top row), versus a wider prior ($N(0,5)$, middle row), versus our new APPS method (bottom row). “X” indicates that a significant confirmation bias was not detected ($p > .05$), while “✓” indicates a significant bias was detected. The tables show findings for both absolute (a) and normalized (b) confirmation bias.

1.8. Confirmation bias overestimation varies as a function of the generative confirmation bias

To better understand how learning rate asymmetries are recovered using MAP estimation, we conducted a sweep analysis across various levels of generative confirmation bias. For each step of the sweep, we fixed the confirmation bias and calculated the learning rates such that the average learning rate remained 0.4 ($\alpha_c = 0.4 + CB/2$, $\alpha_d = 0.4 - CB/2$). The other three parameters (β , τ and ϕ) were held fixed at the mean of the best-fit estimates from the original hybrid model MAP fits of experiment P2. For each level of confirmation bias, we generated 1,001 datasets for the P2 task, fitted them with the full hybrid model (MAP estimation with a ϕ prior $N(0,1)$), and examined the resulting confirmation bias estimates.

In alignment with our main results (Fig. 3c), this analysis revealed a consistent over-estimation of confirmation bias across all examined levels. However, this overestimation was not constant; instead, it followed a non-linear pattern, being especially high for slightly negative generative

values and decreasing as the confirmation bias moved toward more extreme positive or negative values (Supplementary Fig. 9a). This effect was driven by an overestimation of the confirmatory learning rate (α_c) and an even more pronounced underestimation of the disconfirmatory learning rate (α_d) (Supplementary Fig. 9b-c). In Supplementary Note 2.3 we note how this pattern of recovery errors differs when using an MLE fitting procedure.



Supplementary Figure 9: Recovery of confirmation bias parameters under MAP estimation. The plots illustrate parameter recovery from a sweep across different generative levels of confirmation bias ($\alpha_c - \alpha_d$) using the MAP procedure with an $N(0,1)$ prior on ϕ . **(a)** Recovered absolute confirmation bias vs. generative bias. The recovered bias (violet line) is systematically higher than the ground truth (black diagonal line), showing consistent inflation particularly around the zero-bias point (dashed vertical line). **(b)** Recovery of the confirmatory learning rate (α_c). Recovered values (green line) show a consistent overestimation relative to the ground truth. **(c)** Recovery of the disconfirmatory learning rate (α_d). Recovered values (red line) are systematically lower than the generative ground truth, a key driver of the overall bias inflation.

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

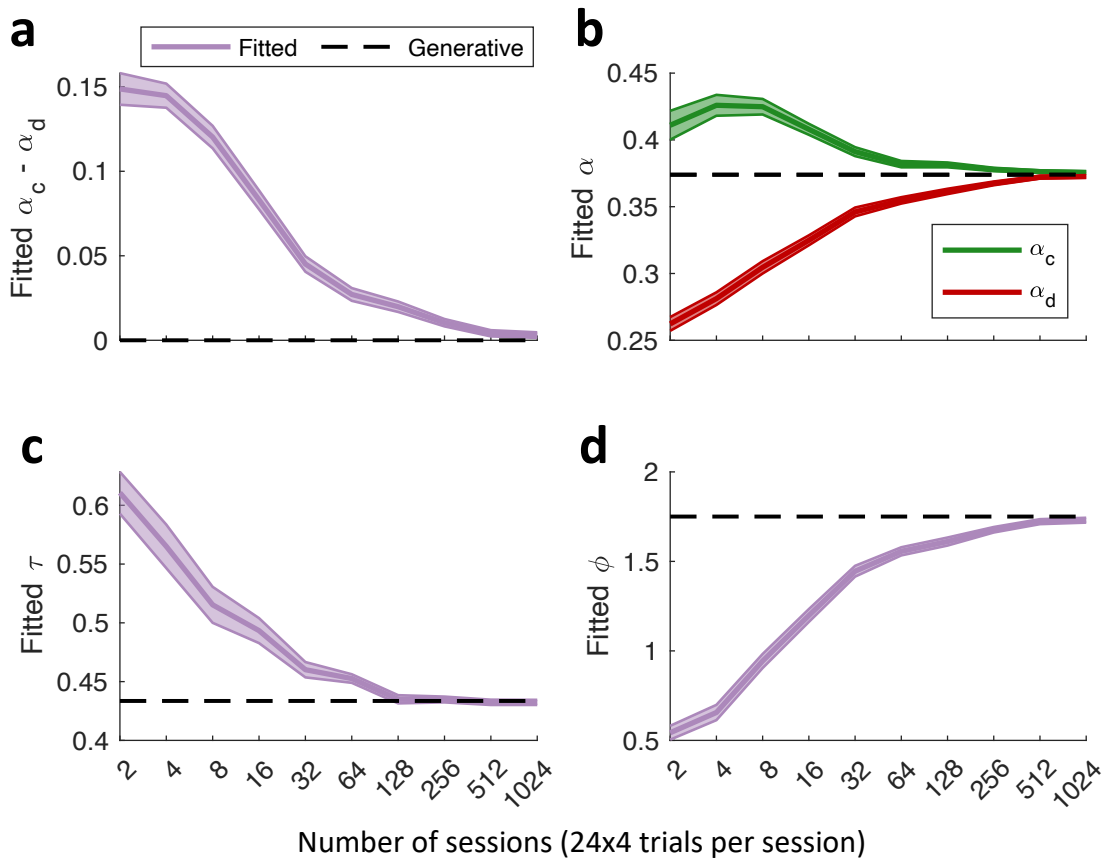
Here, we replicate the asymptotic analysis from the main text (Figure 4), but using the MAP estimation procedure to investigate whether the use of parameter priors exacerbates the small-sample bias. To do this, we first simulated PSL agents using the average empirical parameters obtained from fitting 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 (τ, ϕ) 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 (Supplementary Fig. 10a).

This artifact was driven by a systematic overestimation of α_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 (Supplementary Fig. 10b).

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 (Supplementary Fig. 10c). 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) (Supplementary Fig. 10d).

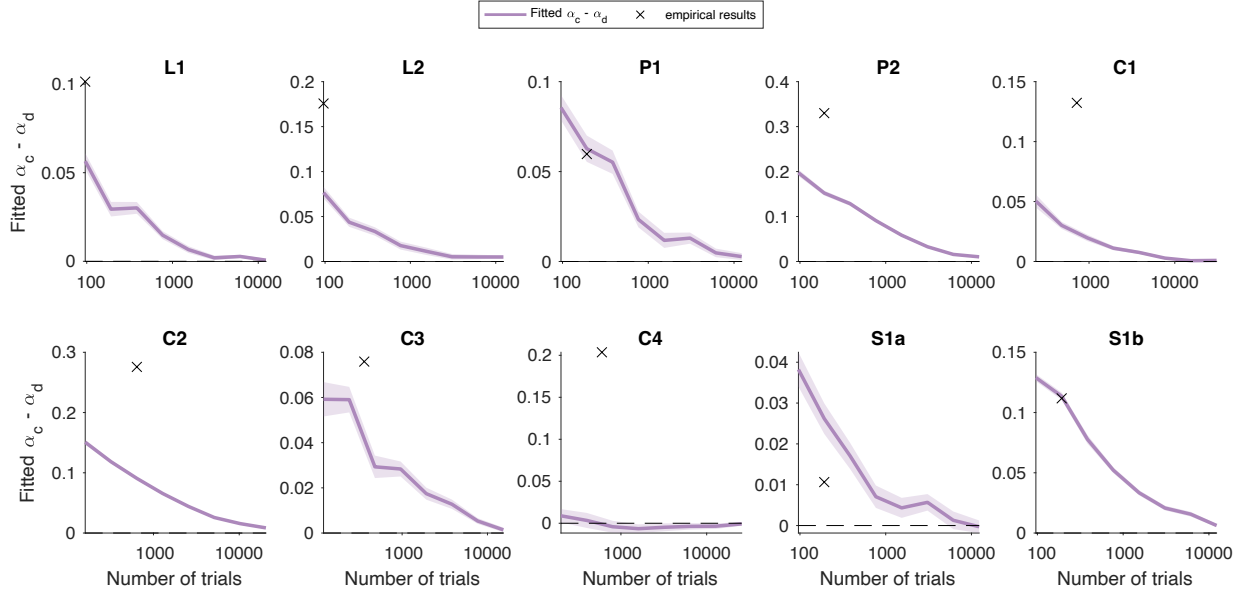
These results highlight the profound impact of parameter priors on estimation. The strong shrinkage effect on the ϕ parameter requires a massive amount of data to overcome, explaining why the MAP procedure needs many more trials to achieve an unbiased dissociation of the mechanisms compared to the MLE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



Supplementary Figure 10: 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 $\pm$ 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.

The above analysis focuses on a particular design and on a single parameter combination (i.e., the group-averaged parameters). However, the magnitude and trajectory of spurious confirmation bias can vary substantially depending on the specific task design and the underlying parameter distribution of a study's sample population. We sought a more detailed impression regarding how spurious learning asymmetries might manifest for cohorts of PSL agents, whose parameters are set to empirical individual parameters, performing each of the experimental designs, as a function of number of sessions. For each of the 10 analyzed designs, and for each examined number of sessions (i.e., 1, 2, 4, 8, 16, 32, 64 and 128), we generated 101 synthetic cohorts (each cohort consisting of one synthetic agent per participant) using the exact task designs and the individual-level empirical parameters (i.e., the empirical PSL parameters fitted with MAP). For each cohort, we simulated data for each agent (for the currently examined number of sessions), refitted these data with the hybrid model, and calculated the group-average spurious learning-asymmetry. Supplementary Figure 11 plots the resulting group-level spurious confirmation bias as a function of the total trial count (x-axes) and design (panel). As illustrated, the resulting curves vary greatly across studies (note the different y-axes scale for different panels): the number of trials required for the spurious bias to dissipate ranges from 200 trials in dataset C4 to well over 10,000 trials in dataset C2. For comparison, we also overlay the empirical group-level confirmation bias (fitted with MAP) observed in the original experiments at their actual trial counts, which, for most studies, sits above the mean spurious baseline predicted for that sample size. However, as we explain across this work, researchers should use the tailored APPS procedure proposed to rigorously confirm that an empirical effect is significant above and beyond the bias expected from a population with overall symmetric learning given potential parameter-fitting biases.



Supplementary Figure 11: Dataset-specific asymptotic reduction of spurious confirmation bias across the ten re-analysed studies using MAP. The plots show the fitted group-level confirmation bias ($\alpha_c - \alpha_d$) as a function of the total number of trials (note the logarithmic scale on the x axes), generated by tailoring simulations to the exact task designs and individual empirical parameter distributions of each of the ten studies. For each study we simulated 1,2,4,8,16,32,64 and 128 sessions (number of trials per session: L1-2/P1-2/S1=96, C1=240, C2=260, C3=120, C4=200). For each study, synthetic datasets were generated by simulating cohorts of PSL agents (based on individual-level empirical parameters), fitting these the hybrid model and calculating the group-average fitted learning asymmetry. Solid purple lines and shaded areas represent the mean $\pm$ SEM across 101 independent cohort simulations per dataset length. The black cross (x) in each panel marks the empirical group-level confirmation bias estimated from the original study's sample at its actual trial count baseline. Dashed horizontal lines indicate a true null confirmation bias ($\alpha_c - \alpha_d = 0$). The trajectories illustrate that the rate and shape of the asymptotic parameter dissociation depend heavily on design features and population characteristics, with the trial count required for the spurious bias to dissipate varying widely across experiments (note the different y-scales in different panels).

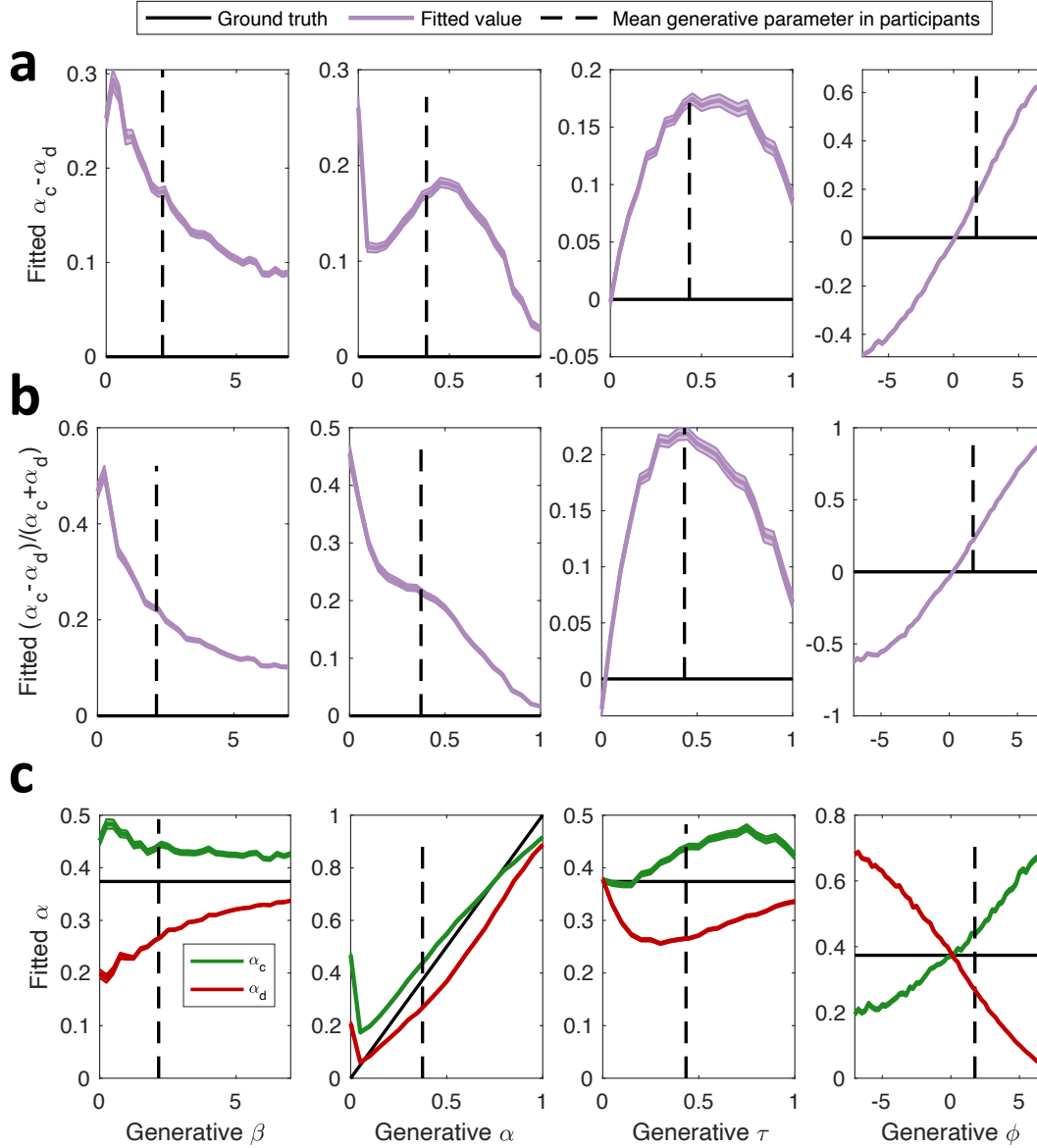
1.10. Contribution of different parameters to the spurious confirmation bias

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) (Supplementary Fig. 12a-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 (Supplementary Fig. 12a-b, middle-left panels). The choice-trace accumulation rate (τ) also exhibited a non-monotonic effect, where its intermediate values produced the largest artifact (Supplementary Fig. 12a-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 ($\phi < 0$) (Supplementary Fig. 12a-c, right panels).

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.



Supplementary Figure 12: 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 $\pm$ 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.

Supplementary note 2: MLE results

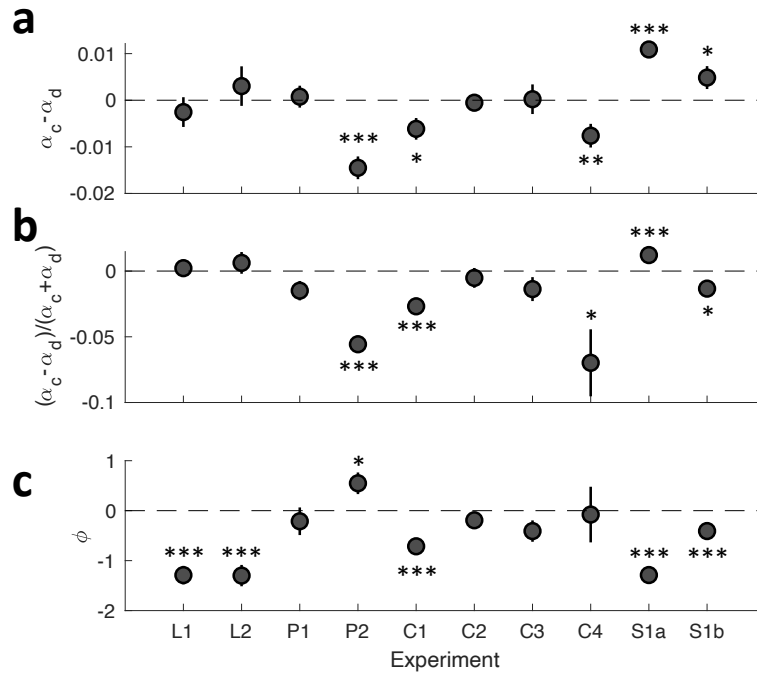
2.1 MLE Does Not Introduce consistent spurious Asymmetries in the Absence of Perseveration

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 Supplementary Note 1.5 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 with those from the MAP procedure. While we still found tiny 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 Supplementary Fig. 13a-b) and sometimes negative (absolute, $p < 0.05$ in P2,C1,C4; normalized, $p < 0.05$ in P2,C1,C4,S1b Fig. ϕ a-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 Supplementary Note 1.5). 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.

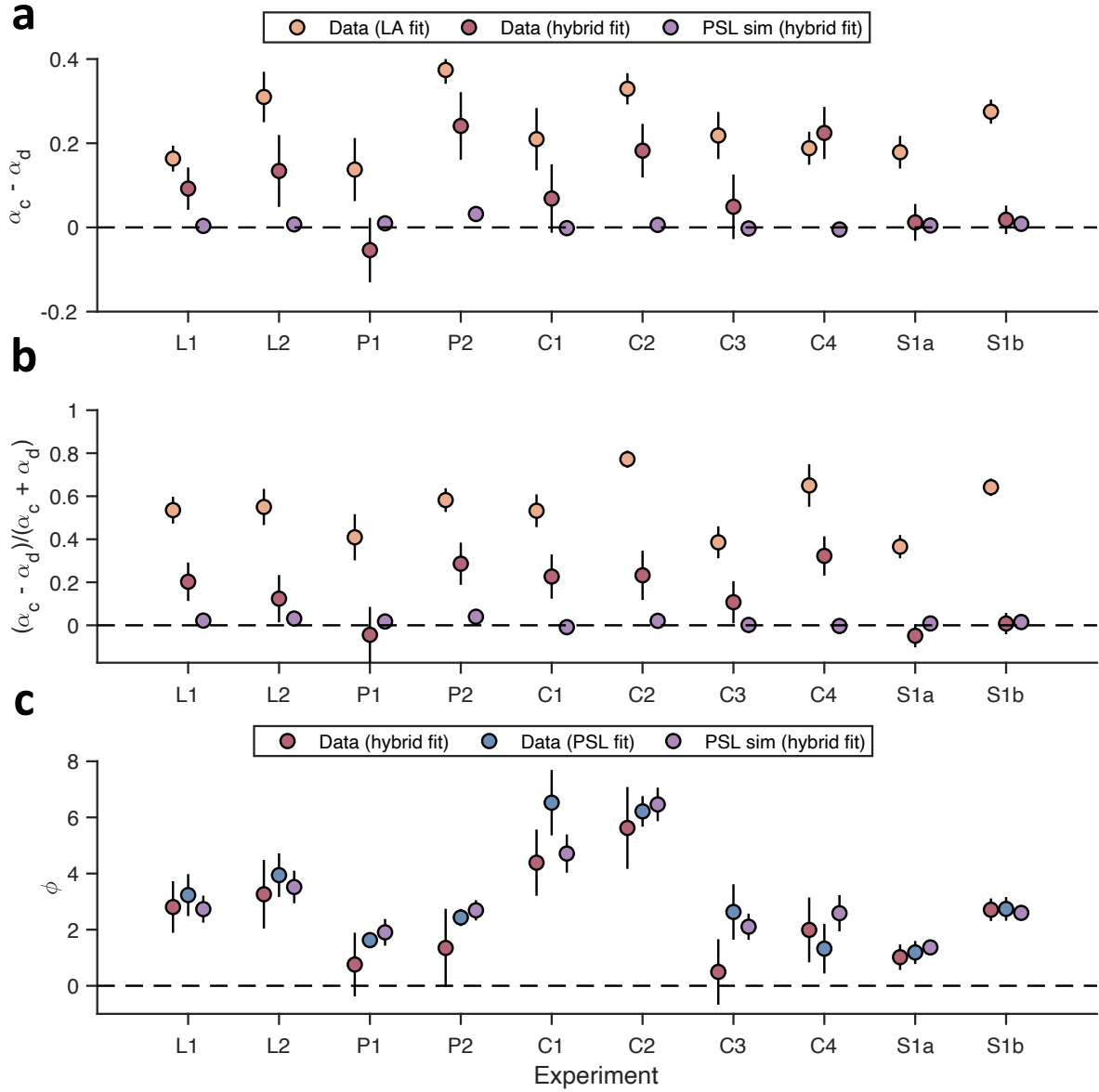


Supplementary Figure 13: Reduced estimation bias in simulations using the MLE procedure. The figure shows the results of fitting the Hybrid model with MLE to simulated data from a simple model with symmetric learning and no perseveration. **(a)** The fitted absolute confirmation bias is not consistently 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 $\pm$ 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 LA 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 L1, 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. 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). Finally, the estimated perseveration weight (ϕ) was not significantly greater for the PSL model than for the hybrid model (Supplementary Fig. 14a-b, red vs. blue points; all p 's > 0.05). Complementing this, our simulation analysis showed that when fitting the PSL simulations with the hybrid model, the recovered ϕ

parameter was not significantly underestimated compared to its true generative value (Supplementary Fig. 14c, violet vs. blue points; all p 's > 0.05 except for C1).



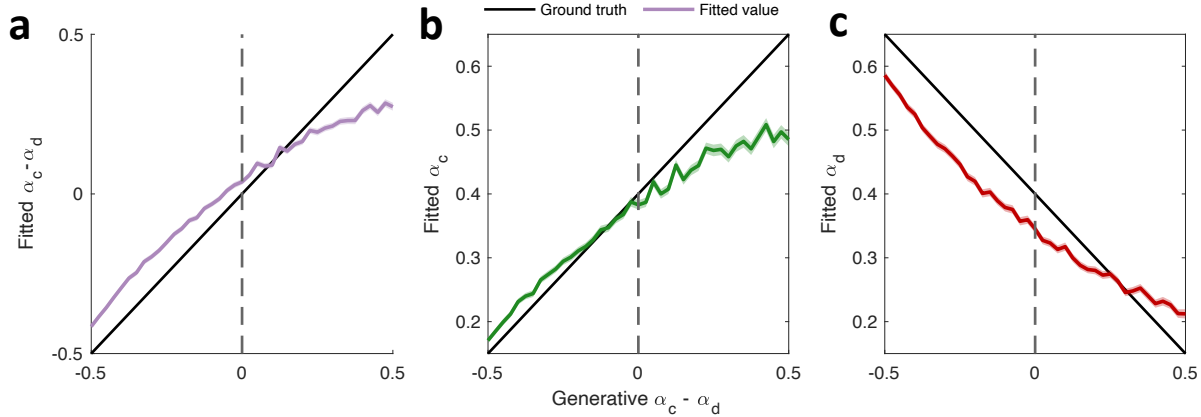
Supplementary Figure 14: Confirmation bias and perseveration parameters estimated from empirical data and simulations fitted with MLE. (a) Absolute confirmation bias ($\alpha_c - \alpha_d$) estimated from fitting participant data with the Pure LA model (yellow points) and the Hybrid model (red points). 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 $\pm$ SEM across participants/simulations.

2.3 Confirmation bias overestimation varies as a function of the generative confirmation bias

To better understand how learning rate asymmetries are recovered using MLE, we conducted a sweep analysis across various levels of generative confirmation bias. For each step of the sweep, we fixed the confirmation bias and calculated the learning rates such that the average learning rate remained 0.4 ($\alpha_c = 0.4 + CB/2$, $\alpha_d = 0.4 - CB/2$). The other three parameters (β , τ and ϕ) were held fixed at the mean of the best-fit estimates from the original hybrid model MAP fits of experiment P2. For each level of confirmation bias, we generated 1,001 datasets for the P2 task, fitted them with the full hybrid model (MLE fitting procedure), and examined the resulting confirmation bias estimates.

This analysis revealed a more complex pattern of recovery errors than the consistent overestimation observed when using the MAP procedure (see Supplementary Note 1.8). Specifically, we found an inflation of confirmation bias for generative levels below 0.1, and an underestimation of the bias for generative levels above this threshold (Supplementary Fig. 15a). This effect was driven by an overall underestimation of the confirmatory learning rate (α_c) for levels above -0.1 (and an overestimation for levels above), combined with an underestimation of the disconfirmatory learning rate (α_d) for generative values below 0.3 (and an overestimation for levels above) (Supplementary Fig. 15b-c).

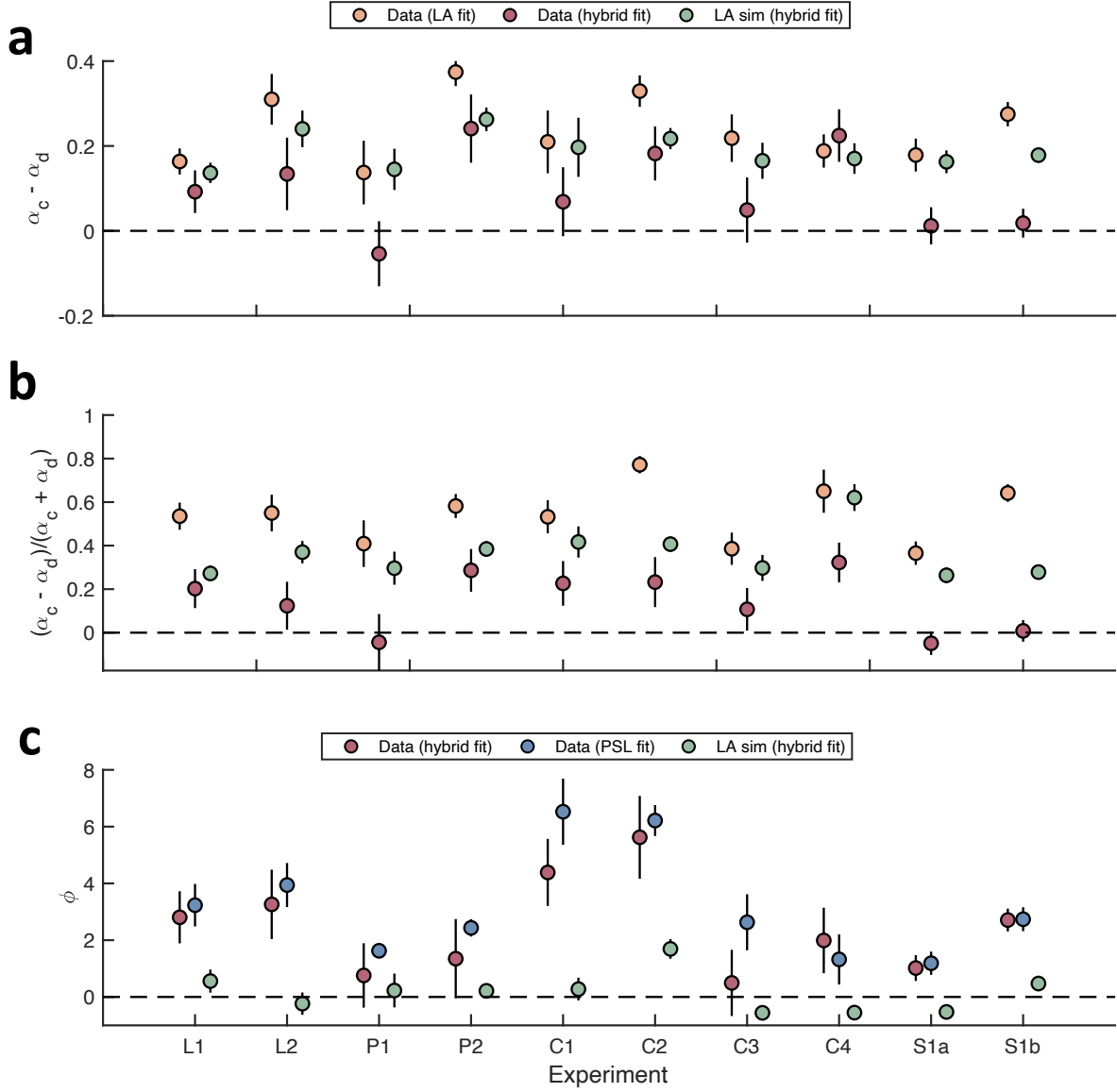


Supplementary Figure 15: Recovery of confirmation bias parameters under MLE estimation. The plots illustrate parameter recovery from a sweep across different generative levels of confirmation bias ($\alpha_c - \alpha_d$) using the MLE procedure. **(a)** Recovered absolute confirmation bias vs. generative bias. Unlike the MAP procedure, MLE shows a non-linear error pattern: the bias is overestimated for generative values below 0.1 and underestimated for values above 0.1. **(b)** Recovery of the confirmatory learning rate (α_c). Recovered values (green line) show an underestimation for levels below -0.1, and an overestimation for levels above this threshold. **(c)** Recovery of the disconfirmatory learning rate (α_d). Recovered values (red line) show a crossover effect, being underestimated for generative values below 0.3 and overestimated for values above 0.3.

2.4 Confirmation bias produces little spurious positive perseveration

As we did for MAP (see Supplementary Note 1.6), we investigated whether behavior generated by a pure learning asymmetries could lead to the spurious detection of perseveration when using MLE. To this aim, we simulated 1001 datasets per participant from our LA model (using the MLE-estimated empirical parameters) and subsequently fitted these datasets with our hybrid model.

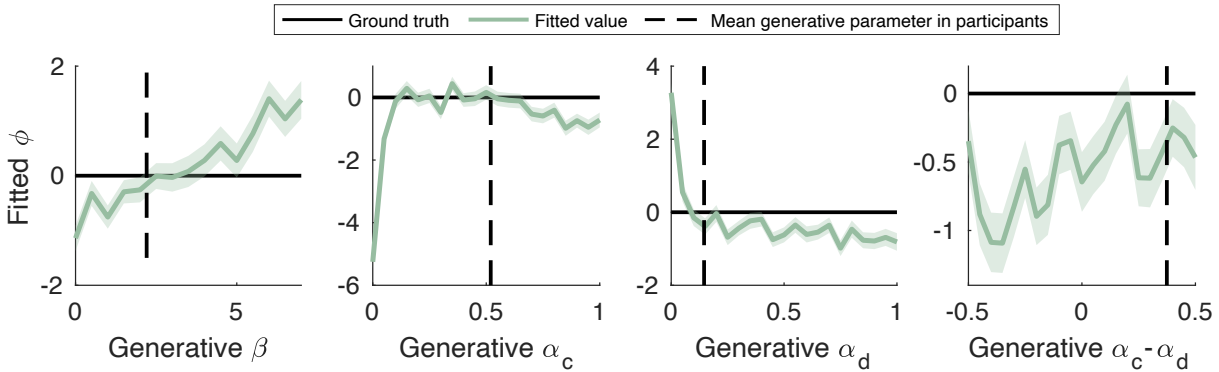
The fitted confirmation bias was underestimated in 4 out of 10 experiments when measured in absolute terms (yellow vs green points in Supplementary Fig. 16a; L2,P2,C2,S1b, $p > 0.05$) and in 8 out of 10 experiments when measured in normalized terms (Supplementary Fig. 16b; all p 's < 0.05 , except for P1 and C4). Despite this tendency, we did not find a consistently positive spurious perseveration. The fitted perseveration parameters (ϕ) were significantly *negative* for 3 experiments (C3,C4 and S1a; all p 's < 0.05), non-significant for five experiments (L1,L2,P1,P2,C1), and only significantly positive for two experiment (C2 and S1a) (Supplementary Fig. 16c). Together, these results suggest that while confirmation bias may be underestimated, the perseveration parameter is not consistently overestimated.



Supplementary Figure 16: Confirmation bias and perseveration parameters estimated from empirical data and LA simulations fitted using MLE. (a) Absolute confirmation bias estimated from fitting participant data with the Pure LA model (yellow points) and the Hybrid model (red points). Additionally, green points show the confirmation bias recovered when fitting LA 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 LA agents, showing that no spurious positive perseveration parameter emerges for most experiments. Points and error bars represent mean $\pm$ SEM across participants/simulations.

We next conducted a parameter sweep analysis following a similar procedure as in Supplementary Note 1.10. To do this, we systematically varied each of the three key generative parameters of the Pure Learning Asymmetry (LA) model (β , α_c , α_d) and the net confirmation bias ($\alpha_c - \alpha_d$) across a range of values. For each step of the sweep, one parameter was varied across a grid of values, while the other parameters were held fixed at the mean of the best-fit estimates from the original LA MLE 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 recovered perseveration parameter (ϕ). This allowed us to characterize which parameters of the LA model most strongly drive the emergence of artifactual perseveration.

This analysis revealed complex relationships between the generative parameters and artifactual perseveration. First, when the inverse temperature (β) as low, we detected a negative ϕ value, which increased as the inverse temperature increased, eventually reaching positive values (Supplementary Fig. 17, left panel). The true learning rates (α_c , α_d) showed mirrored patterns: low values of α_c produced high negative perseveration, while low values of α_d produced highly positive perseveration. As these learning rates increased, the spurious perseveration asymptoted to slight negative values, especially for the disconfirmatory learning rate (Supplementary Fig. 17, middle panels). Finally, we detected a negative spurious perseveration for all levels of confirmation bias, with a slight positive trend as the confirmation bias magnitude increased (Supplementary Fig. 17, right panel). This shows how biases in estimating perseveration in the presence of learning asymmetries may still emerge and are highly sensitive to parameter values.

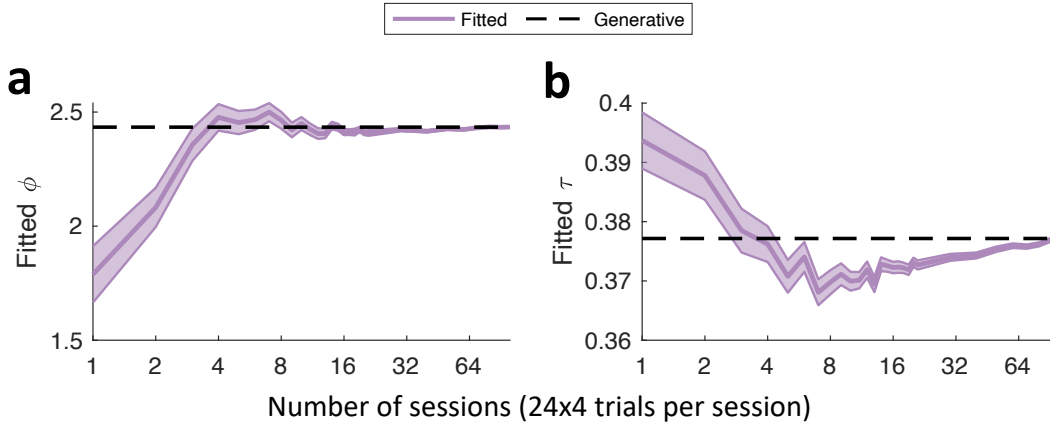


Supplementary Figure 17: Parameter sweep showing how generative parameters from the LA model (asymmetric learning, without perseveration) influence spurious perseveration detection. The figure shows the results of a parameter sweep designed to isolate the influence of each generative parameter (β , α_c , α_d , and $\alpha_c - \alpha_d$). 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 LA model fits of experiment P2.

2.5 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 (Supplementary Fig. 18a). 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 (Supplementary Fig. 18b).

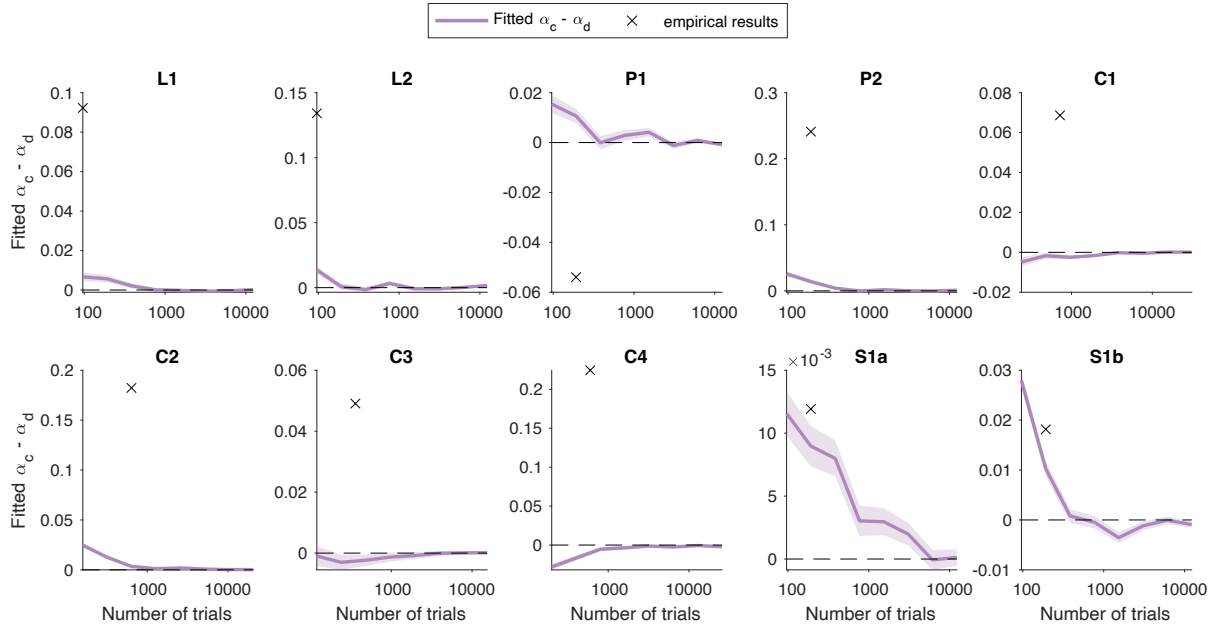


Supplementary Figure 18: Misestimation of perseveration parameters fitted with MLE is reduced with dataset size. The figure shows parameter estimates obtained by fitting the Hybrid model to simulations of PSL agents, as a function of the number of sessions in the simulated dataset (where each session 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 $\pm$ 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.

The above analysis focuses on a particular design and on a single parameter combination (i.e., the group-averaged parameters). However, the magnitude and trajectory of spurious confirmation bias can vary substantially depending on the specific task design and the underlying parameter distribution of a study's sample population. We sought a more detailed impression regarding how spurious learning asymmetries might manifest for cohorts of PSL agents, whose parameters are set to empirical individual parameters, performing each of the experimental designs, as a function of number of sessions. For each of the 10 analyzed designs, and for each examined number of sessions (i.e., 1, 2, 4, 8, 16, 32, 64 and 128), we generated 401 synthetic cohorts (each cohort consisting of one synthetic agent per participant) using the exact task designs and the individual-level empirical parameters (i.e., the empirical PSL parameters fitted with MLE).

Supplementary Figure 19 plots the resulting group-level spurious confirmation bias under MLE as a function of the total trial count (x-axes) and design (panel). As illustrated, spurious biases using

MLE are weaker than with MAP (compare with Supplementary Figure 11). Yet, the resulting spurious curves vary across studies (note the different y-axes scale for different panels). For instance, some studies actually predict a small negative spurious confirmation bias (i.e., a spurious disconfirmation bias) at low trial counts (see C1 and C4). Despite these diverse trajectories, all analyzed studies eventually show the expected asymptotic spurious-bias dissipation, though the of number of trials needed to reach this asymptotic regime is highly study-dependent (ranging from fewer than 200 trials in dataset C4 to nearly 10,000 trials in dataset S1a). For comparison, we also overlay the empirical group-level confirmation bias (fitted with MLE) observed in the original experiments at their actual trial counts, which, for most studies, sits above the mean spurious baseline predicted for that sample size (note such spurious baselines are even greater when using MAP; see Supplementary Note 1.9, Supplementary Fig. 11). However, as we explain across this work, researchers should use the tailored APPS procedure proposed to rigorously confirm that an empirical effect is significant above and beyond the bias expected from a population with overall symmetric learning given potential parameter-fitting biases.

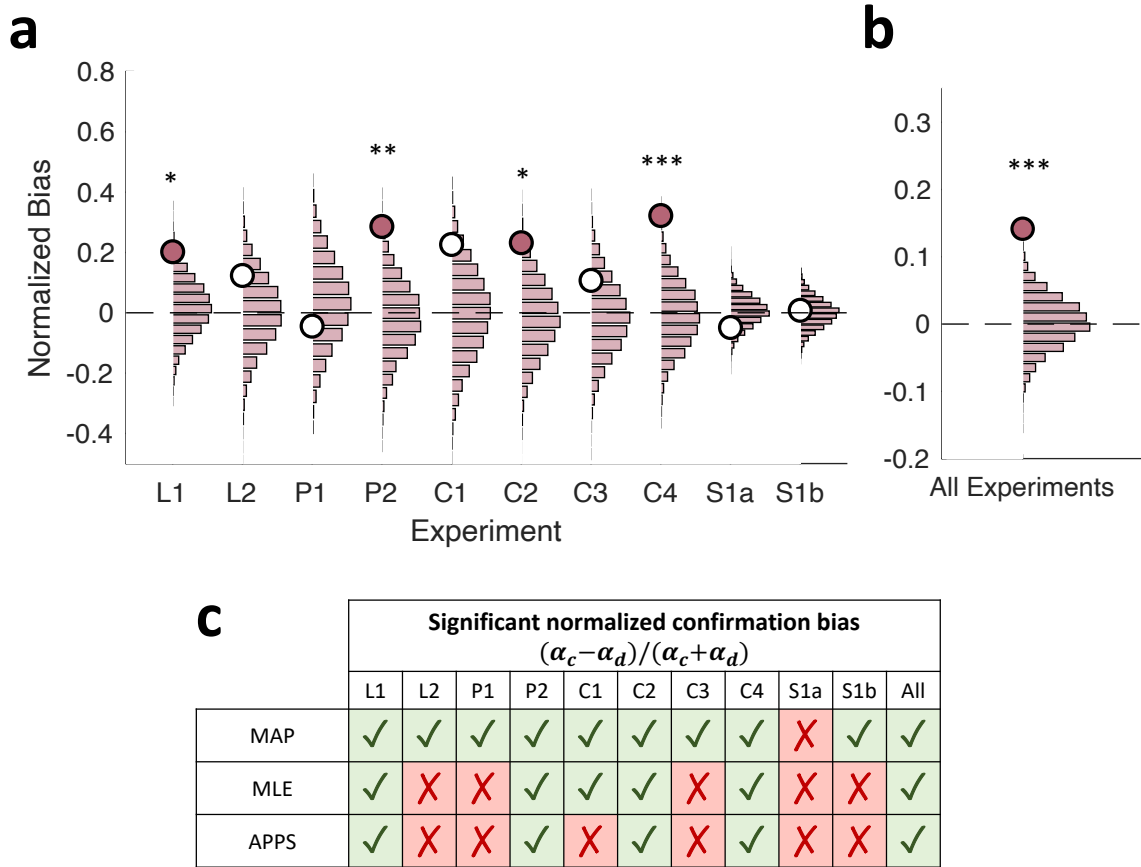


Supplementary Figure 19: Dataset-specific asymptotic reduction of spurious confirmation bias across the ten re-analysed studies using MLE. The plots show the fitted group-level confirmation bias ($\alpha_c - \alpha_d$) as a function of the total number of trials (note the logarithmic scale on the x axes), generated by tailoring simulations to the exact task designs and individual empirical parameter distributions of each of the ten studies. For each study we simulated 1,2,4,8,16,32,64 and 128 sessions (number of trials per session: L1-2/P1-2/S1=96, C1=240, C2=260, C3=120, C4=200) For each study, synthetic datasets were generated by simulating cohorts of PSL agents (based on individual-level empirical parameters), fitting these the hybrid model and calculating the group-average fitted learning asymmetry. Solid purple lines and shaded areas represent the mean $\pm$ SEM across 401 independent cohort simulations per dataset length. The black cross (x) in each panel marks the empirical group-level confirmation bias estimated from the original study's sample at its actual trial count baseline. Dashed horizontal lines indicate a true null confirmation bias ($\alpha_c - \alpha_d = 0$). The trajectories illustrate that the rate and shape of the asymptotic parameter dissociation depend heavily on design features and population characteristics, with the trial

count required for the spurious bias to dissipate varying widely across experiments (note the different y-scales in different panels).

2.6 Group-level test against a null distribution of normalized spurious confirmation bias

We repeated our APPS procedure focusing on the normalized metric of confirmation bias. Again, only 3 out of the 10 experiments (L1, P2, and C4) demonstrated a significant absolute confirmation bias under this stricter test ($p < 0.05$; Supplementary Fig. 20a). APPS changed the conclusion compared with a simple MLE procedure combined with standard t-testing. For example, in experiments C1, a standard fit produces a significant group-level bias ($p < 0.05$; one-tailed t-test), but our APPS approach, we find that this bias is not significantly greater than the artifactual bias expected from a group with overall symmetric learning coupled with fitting biases and perseverative leakage (Supplementary Fig. 20a,c). Finally, a meta-analytic version of our method confirmed a significant and robust normalized confirmation bias at the meta-analytic level ($p < 0.001$; Supplementary Fig. 20b).



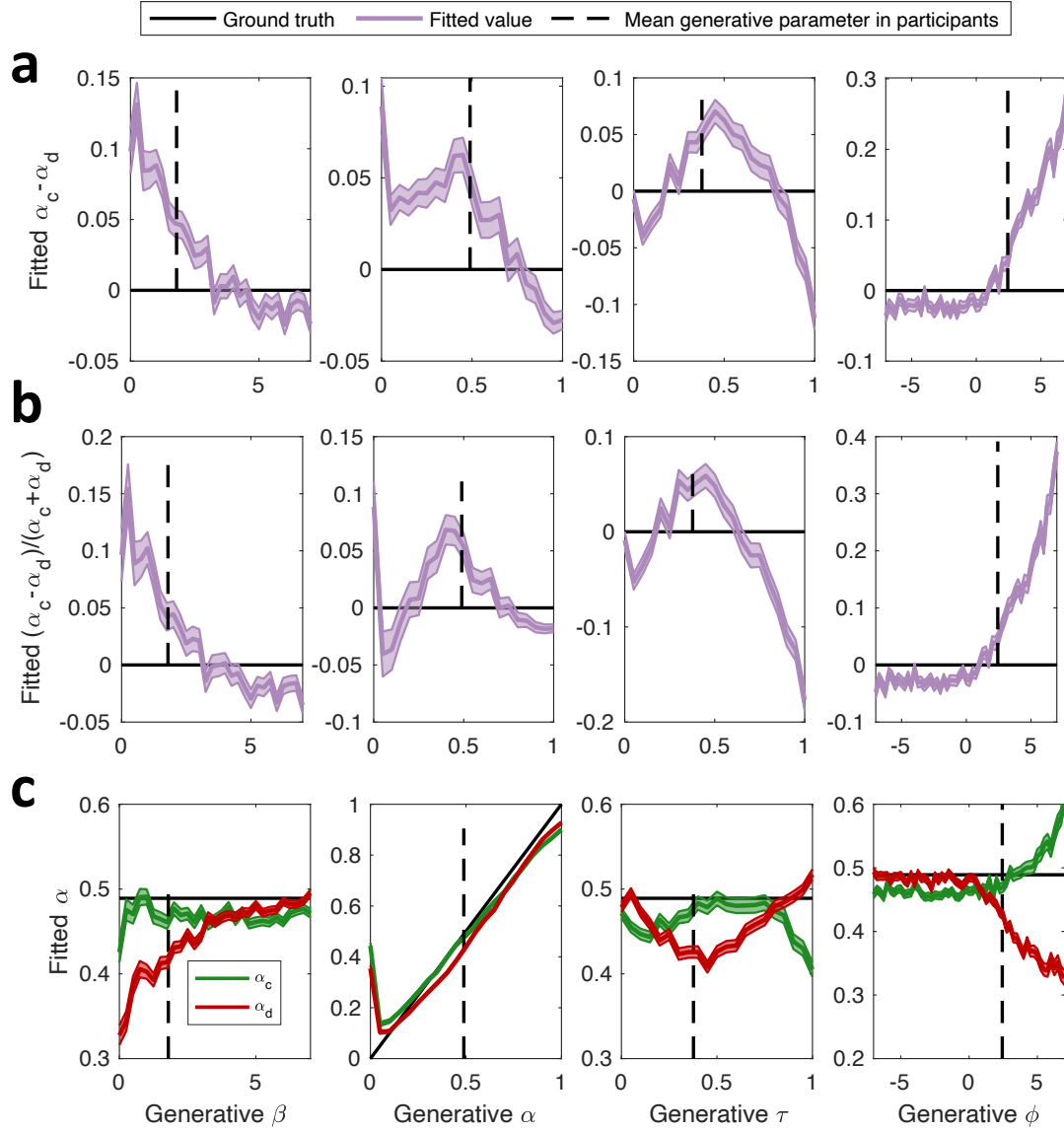
Supplementary Figure 20: Group-level hypothesis testing for a confirmation bias against a null distribution of confirmation bias affected by fitting artifacts and model mimicry. (a) Normalized confirmation bias results across individual experiments. Each histogram represents the null distribution

of the mean artifactual bias generated via the APPS procedure for that specific experiment. The overlying circles represent the mean empirical normalized confirmation bias. Circles are colored red if the empirical mean was significantly larger than the null distribution **(b)** Meta-analytic result. The plot shows the mean absolute confirmation bias aggregated across all 10 experiments (*, $p < 0.05$; **, $p < 0.01$, ***, $p < 0.001$). $((\alpha_c - \alpha_d)/(\alpha_c + \alpha_d) > 0)$ Comparison of qualitative conclusions across fitting and testing methods. The table summarizes the detection of a significant normalized confirmation bias (2) for each experiment across three methods: (i) standard MAP estimation with $N(0,1)$ perseveration prior; (ii) standard MLE estimation; and (iii) the new APPS method. The green checkmarks indicate a significant result ($p < 0.05$), while red "X" marks indicate non-significance.

2.7 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 Supplementary Note 1.10. 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.

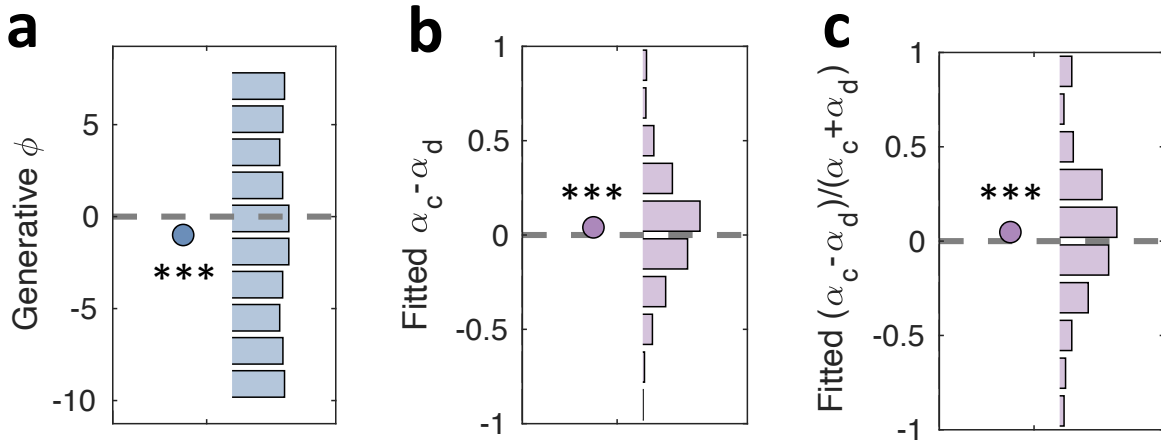
This analysis again revealed complex relationships between the generative parameters and the artifactual bias. The relationship with the generative parameters was broadly the same as that observed when using MAP estimation, but with a few key differences. First, the large bias for low values of the choice 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) (Supplementary Fig. 21a-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 (Supplementary Fig. 21a-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 (Supplementary Fig. 21a-c, right panels).



Supplementary Figure 21: 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 $\pm$ 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.

2.8 A group with average negative perseveration can produce a significant spurious confirmation bias

Crucially, the asymmetry in the effect of ϕ (Figure 4b and Supplementary Figure 21) implies that even a sample of participants with overall anti-perseverative tendencies (i.e., mean $\phi < 0$) can still produce a net spurious confirmation bias at the group level. This is because strong positive artifacts generated by perseverative individuals in the population can overwhelm weaker negative artifacts from the anti-perseverative individuals. To test this explicitly, we generated 2,000 simulations of a PSL population with an average negative ϕ (average $\phi = -1$, $p < 0.001$) (Supplementary Fig. 22a) (see “Methods: Simulation of a Population with Negative Perseveration”). When this dataset was fitted with the hybrid model, we detected a significantly positive confirmation bias, for both the absolute and normalized metrics ($p < 0.001$; Supplementary Fig. 22b-c). These results illustrate that a significant, albeit spurious, confirmation bias can emerge at the group level even when the population’s average tendency is anti-perseverative.



Supplementary Figure 22: A population with a group-level negative perseveration can still produce a spurious confirmation bias. (a) Distribution of perseveration parameters (ϕ) from the PSL simulations, which follows a uniform distribution with a mean of -1. (b) The absolute confirmation bias detected when fitting these PSL simulations using MLE. 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.

Supplementary note 3: New hypothesis-testing procedures

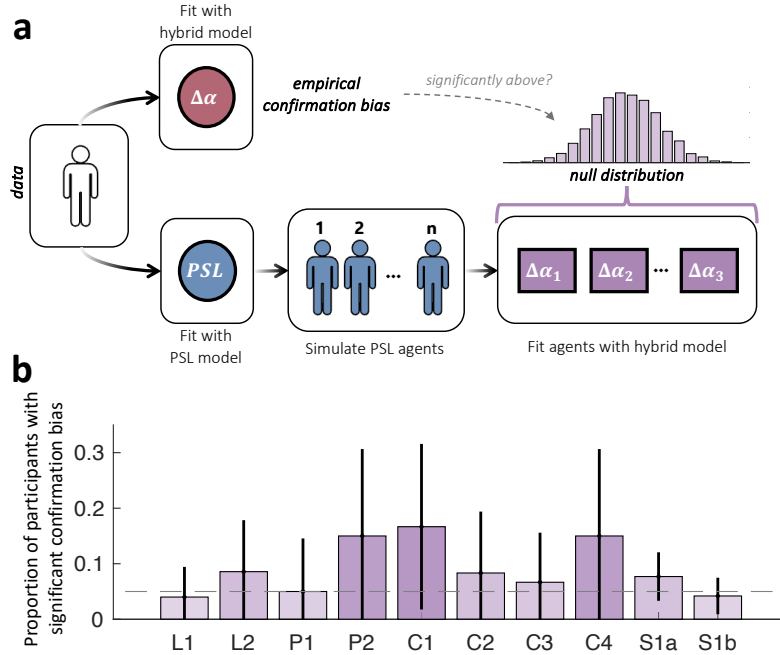
Note: In this section all fits were done based on MLE unless stated otherwise.

3.1 Individual-level test against a null distribution of spurious confirmation bias

3.1.1 Test results

Apart from our group-level test reported in the main text, we propose a new statistical test for the null hypothesis (H0) that learning is *symmetric* (and might exhibit choice perseveration) against the alternative hypothesis (H1) of asymmetric learning + perseveration at an individual level. We describe how we applied this test for our 10 datasets. For each participant we generate, using a parametric-bootstrap approach, a null distribution of the spurious confirmation bias expected under H0 (when fitting the hybrid model), in comparison to which we can determine whether an individual's empirical bias is statistically significant or simply reflects biased parameters. First, to estimate the degree of choice perseveration under the assumption of symmetric learning (H0), we fitted each participant's empirical data with the PSL model to obtain their individual best-fit parameters. Second, using these parameters, we generated 1,001 simulated behavioral datasets from the PSL model for each participant. Third, each of these simulated datasets was fitted with the full hybrid model, yielding a pool of 1,001 "spurious bias" estimates (the difference between confirmatory and disconfirmatory learning rates) per participant (Supplementary Fig. 23, bottom). Finally, we calculated an individual p-value as the proportion of these null samples that were larger than or equal in magnitude to the participant's empirical confirmation bias (obtained from the initial hybrid model fit; Supplementary Fig. 23a, top).

This analysis provides evidence for a significant confirmation bias (above and beyond any spurious bias produced by perseveration) only a minority of participants (Supplementary Fig. 23b). Across all experiments, the proportion of participants showing a significant confirmation bias ($p < 0.05$) ranged from 4% (in L1; i.e., around the type I error rate) to 16.7% (in C1). However, it is important to consider that common RL experiments are better powered to detect significant population-level effects and therefore, these individual-level results may also reflect low statistical power (see Supplementary Note 3.1.2 for a characterization of the test's power).



Supplementary Figure 23: Testing empirical confirmation bias in individual participants against a null distribution of artefactual bias generated from PSL simulations. All model fits used in this procedure are based on MLE. **(a)** Schematic of the individual bootstrapping procedure. For each participant, an empirical confirmation bias is estimated (top) by fitting the participant’s data to the hybrid model. A null distribution of spurious confirmation bias is then generated by fitting the participant’s data with the PSL model (symmetric learning + perseveration), simulating 1,001 PSL agents, and fitting them with to the hybrid (bottom). **(b)** Proportion of participants in each experiment showing a significant confirmation bias ($p < 0.05$) above and beyond the spurious model mimicry confound produced by perseveration. Error bars represent 95% confidence intervals, the dashed horizontal line represents the type I error rate of 5%.

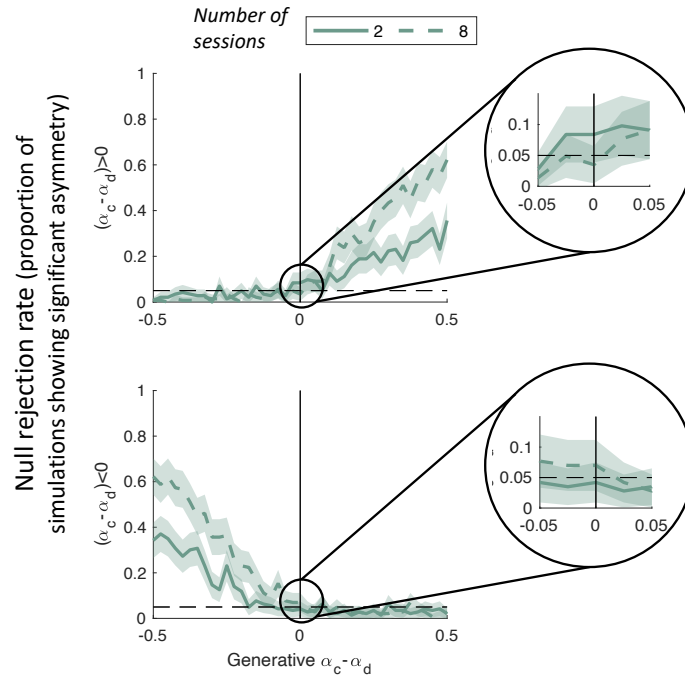
3.1.2 Power analysis

To characterize the statistical power of our individual-level parametric bootstrapping test, we performed a series of simulations using experiment L1 as a representative example. We utilized the best-fitting parameters from the hybrid model for the 50 participants in the experiment. We then systematically varied the learning rates—while keeping all other parameters fixed—to span different levels of generative learning rate asymmetries (ranging from -0.5 to +0.5), while maintaining a constant average learning rate of 0.4 (i.e., $\alpha_c = 0.4 + CB/2$, $\alpha_d = 0.4 - CB/2$).

For each synthetic participant, we generated 3 simulations of the hybrid model (150 simulations in total for each CB level) and fitted them with the hybrid model to obtain an empirical measure of confirmation bias estimates per generative CB level. We then applied our individual-level test to each of these hybrid model simulations, generating 201 PSL simulations and fitting them with the hybrid model to construct a null distribution. The empirical confirmation bias was compared to this null distribution to determine if 1-sided test could detect a significant asymmetry ($p < 0.05$). This was defined as the empirical bias exceeding the 95th percentile of the null for positive confirmation

bias ($\alpha_c - \alpha_d > 0$; Supplementary Fig. 24 top), or falling below the 5th percentile for negative confirmation bias ($\alpha_c - \alpha_d < 0$; Supplementary Fig. 24 bottom). We repeated this procedure for two experiment lengths: 2 sessions (approximately 200 trials, reflecting the original experiment) and 8 sessions (approximately 800 trials).

The results, presented in Supplementary Figure 24, show the proportion of participants displaying a significant confirmation bias above and beyond perseveration for each level of generative bias. Crucially, when the generative confirmation bias is zero (representing the null hypothesis), the positive rate (Type I error) remains at approximately 0.05 (Supplementary Fig. 24 zoomed-in panels). This indicates that the bootstrapping procedure provides adequate control for type-I error rates. However, the analysis also reveals that the individual-level test is relatively underpowered. For 2 sessions (the most common length in our aggregated dataset), power reached approximately 40% only at high generative values (± 0.5). For the average bias level observed empirically in participants ($+0.1$), the power was notably low, between 10–15%. Increasing the trial count to 800 (8 sessions) improved the power, but only to approximately 60% at high values, indicating that it is difficult to achieve high power to detect individual-level effects using this design with a feasible number of trials. We note that our finding that only about 4–16% of participants across experiments show a significant confirmation bias (Supplementary Fig. 23b) is likely driven by this limited statistical power.



Supplementary Figure 24: Power analysis of the individual-level bootstrapping test for different levels of confirmation bias. The plots show the proportion of synthetic agents identified as having a significant learning asymmetry (y-axis) as a function of the generative learning rate asymmetry (x-axis), using our new statistical t-test (1-sided). The top panel shows results for a statistical test for positive confirmation bias ($\alpha_c > \alpha_d$) and the bottom panel for negative (disconfirmation) bias ($\alpha_c < \alpha_d$). Solid lines represent results for 2 sessions (200 trials); dashed lines represent results for 8 sessions (800 trials). Shaded areas represent the 95% confidence interval.

3.2 Group-level test against a null distribution of spurious confirmation bias (APPS Test)

3.2.1 Power Analysis

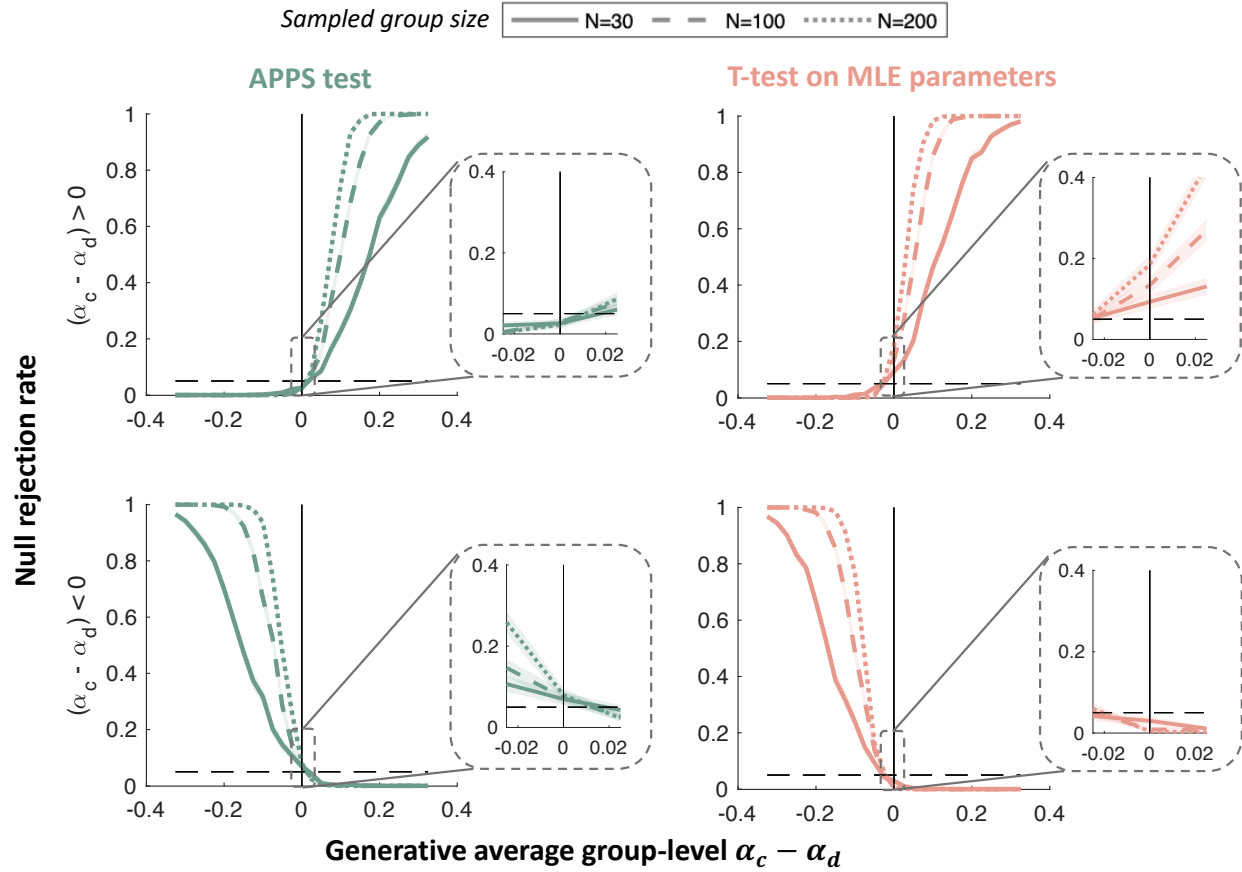
We examined the statistical power of the Anti-Participant Permutation Simulation (APPS) test to detect significant group-level learning asymmetries both as a function of the population-level asymmetry and sample size ($N_{SZ} = 30, 100$, and 200 participants). We utilized the design of experiment L1 ($N=50$) as an illustrative baseline for these simulations. For each target sample size (N_{SZ}) and generative population-level learning asymmetry (m), we performed the following procedure:

- 1) We sampled N_{SZ} individuals (with replacement) from the original participant pool. For each individual, we sampled a confirmation bias (CB) from a normal distribution centered at m with a standard deviation of 0.1 ($N(m, 0.1^2)$), truncated at two standard deviations. We then created a synthetic hybrid agent using the empirical best-fitting parameters of the corresponding participant, with the exception of the learning rates. Specifically, the confirmatory learning rate (α_c) was sampled from a uniform distribution bounded to allow for the specified confirmation bias level (e.g., $\alpha_c \sim U(0.2, 1)$ for $CB = 0.2$, or $\alpha_c \sim U(0, 0.8)$ for a $CB = -0.2$) and the disconfirmatory learning rate was simply fixed to $\alpha_d = \alpha_c - CB$.
- 2) We simulated synthetic data for each of the N_{SZ} agents in our sample, fitted these simulated data with the hybrid model and calculated a group average learning asymmetry effect.
- 3) We applied the full APPS procedure (as described in the main text and Figure 5) to the resulting group data. For comparison, we also performed a standard one-tailed t-test on individually estimated MLE parameters from step 2. In both cases, we checked whether the null hypothesis was rejected at an alpha level of 0.05 .

The above procedure was repeated $1,001$ times per parameter combination (each level of m and N_{SZ}) to calculate the null rejection rate: the proportion of simulated groups in which a significant positive (Supplementary Fig. 24, top) or negative (Fig. CB , bottom) asymmetry was detected. Note that when the generative mean $CB = 0$, this rejection rate represents the type I error rate.

For levels of bias typically observed in empirical reinforcement learning datasets (around 0.1), a standard t-test on MLE parameters demonstrated high power, reaching 100% for groups of 200 and 85% for groups of 100 participants (Supplementary Fig. 24, top right). However, this seemingly high power was accompanied by a grossly inflated Type I error rate. When the generative group-level confirmation bias was zero ($CB = 0$), the standard t-test yielded false positives in approximately 11% , 15% , and 20% of cases for groups of 30 , 100 , and 200 participants, respectively (Supplementary Fig. 24, top right). By contrast, the APPS method maintained robust Type I error control, with rejection rates remaining centered around the nominal 5% level across all group sizes when $CB = 0$ (Supplementary Fig. 24, top left). The APPS test approached reasonable power for the large sample size (75% for $N=200$ participants).

For completeness, we also evaluated the power of these tests to negative learning asymmetries (disconfirmation bias; bottom panels). Here, the APPS test demonstrated superior power compared to the standard MLE-based t-test, (however, note a slightly inflated, Type I error rate of ~7% for all group sizes). For example, at a generative disconfirmation bias of $CB = -0.1$ and a sample size of 200, the APPS test reached a power of 94%, whereas the standard t-test only reached 73%. Here, MLE's power seemed to be compromised by being too conservative (type I error rate ~1%).



Supplementary Figure 24: Power analysis of APPS method and standard t-test on MLE parameters for different levels of confirmation bias. The plots show the probability of detecting a significant group-level learning-asymmetry effect (positive in top panel, negative in bottom panel) as a function of the population level learning asymmetry (x-axes), and simulated sample-size (30 in solid line, 100 in dashed line, 200 in dotted line), and the test applied (APPS in green, t-test on MLE parameters in peach). Shaded areas represent 95% intervals on the proportion.

3.2.2 Comparison of type I error rates across methods

We next investigated how well the different statistical tests examined here control for false positives (Type I error rates) at the group level across the 10 datasets. To achieve this, we generated a synthetic population for each experiment that combined the hybrid-model parameters of both the original participants and their corresponding anti-participants (with their confirmatory

and disconfirmatory learning rates swapped). If an original experiment had N participants, our synthetic population consisted of $2N$ agents. By design, this agent-population had a net learning asymmetry bias of exactly zero (i.e., every participant's bias was canceled out by an anti-participant's mirrored bias) while preserving the same distribution of other parameters (such as choice stochasticity or perseveration) as the empirical sample.

We applied our APPS test to agent-samples drawn from this balanced agent-population to estimate the Type I error rate. For a single test of the null hypothesis, we conducted the following steps:

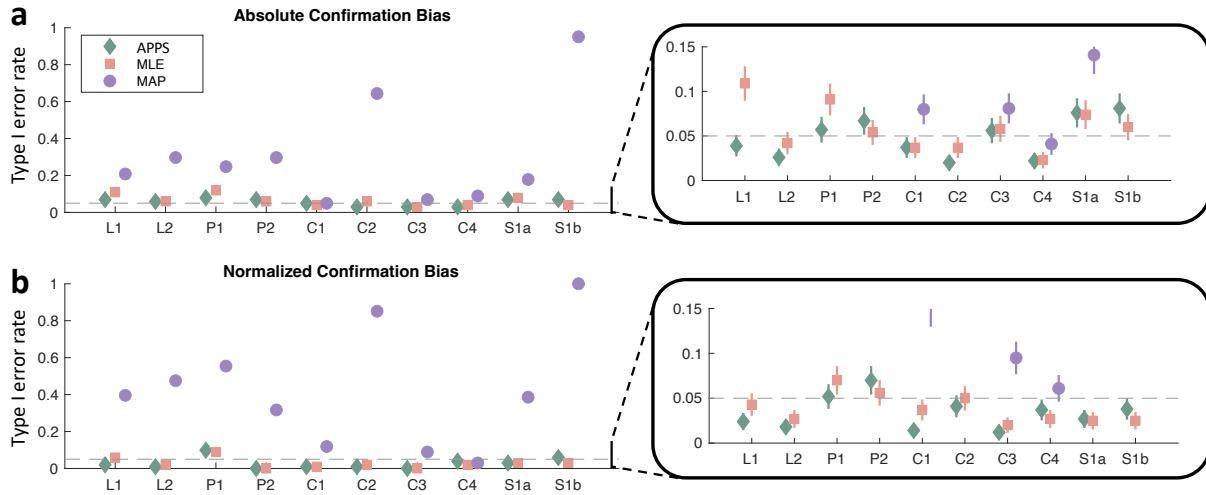
- 1) We sampled (with replacement) N agents from the $2N$ agent population. We refer to these sampled agents as *pseudo-participants* to distinguish them from the empirical participants.
- 2) For each pseudo-participant, we generated a behavioural pseudo-dataset by forward-simulating the experimental task. We then fitted the hybrid model to each pseudo-dataset and averaged the resulting learning-asymmetry estimates across the N pseudo-participants to obtain a simulated group-level pseudo-empirical effect under the null.
- 3) Next, we applied the full APPS statistical procedure to these simulated pseudo-data (as if they were real empirical data). In effect, we created for each pseudo-participant a pseudo agent (an instance of the hybrid model parameterized by the corresponding pseudo participant's parameters) and pseudo anti-agent (same but with the two learning rates shuffled). To generate the APPS null distribution, we permuted within each pseudo-participant between the pseudo-agent and pseudo-anti-agent 1,001 times. In each permutation iteration, we simulated behavior from the pseudo-agent (or pseudo anti-agent) and fitted this behaviour with the hybrid model to obtain a control learning rate asymmetry. We averaged across the control estimates across the N pseudo agents (or anti agents) in the current iteration to obtain a single value in the null distribution. After 1,001 permutations, we calculated the p-value of the simulated empirical group effect.

Finally, to estimate the Type I error rate, we repeated this entire procedure (Steps 1-3) 1,001 times, and calculated the proportion of times the null was rejected.

Because the procedure described above is computationally intensive, we implemented several simplifications to reduce the total processing cost. Instead of generating behavior de novo for every individual sample, we pre-generated 30 simulated pseudo-datasets for each of the $2N$ agents, creating a pool of $2N \times 30$ total pseudo-datasets and their corresponding hybrid fit parameters (pseudo-empirical effects). When step 2 was executed, we randomly selected one of these pre-pseudo-datasets for each pseudo-participant (sampled in step 1) to calculate (based on the corresponding estimated parameters) the pseudo-empirical group-level effect size. We similarly pre-calculated the corresponding null effects for the APPS tests by simulating and fitting 30 pseudo-agents and 30 pseudo-anti-agents for each pseudo-dataset, resulting in a large pool of $2N \times 30 \times 2 \times 30$ pseudo-null effects. When computing the APPS test on a sample (Step 3), we simply permuted over these pre-calculated values for the N sampled pseudo-participants (each with 2×30 precomputed pseudo-(anti)agent null values) to generate the null distribution. This approach removed the necessity to perform new simulations and model fits for every sampling iteration, significantly increasing computational efficiency.

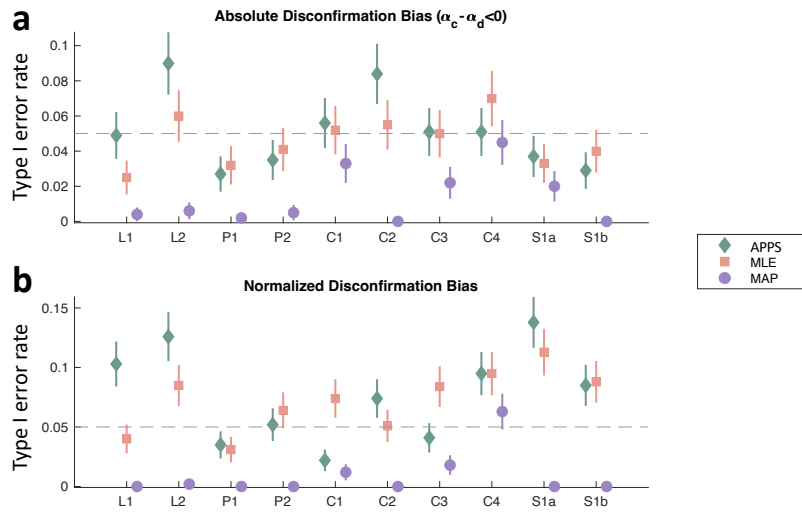
For comparison, we also calculated the Type I error rates for the “standard approaches” MLE and MAP procedures by performing one-tailed t-tests on learning asymmetries estimated from the pseudo datasets on each iteration of step 2 (directly fit with either MLE or MAP with a $N(0,1)$ perseveration prior respectively).

This investigation revealed that the standard MAP procedure produced systematically inflated false-positive rates across experiments, far-exceeding the 5% threshold for both absolute and normalized confirmation bias metrics (Supplementary Fig. 26 violet circles). By contrast, both the MLE (Supplementary Fig. 26 peach squares) and APPS (Supplementary Fig. 26 green diamonds) maintained Type I error rates around the nominal 5% level (but note a minority of cases where type-I error rates seem slightly inflated). Interestingly, type I error rates for MLE and APPS mostly overlapped but in the 2 cases where the methods differed (L1, P1; Supplementary Fig. 26 top), MLE had inflated type-I error rates whereas APPS provided adequate control. This suggests that while individual MLE parameter estimates may be biased, in most of our experiments the method nevertheless provides adequate control for group-level inference within the typical parameter ranges in most of our experiments. In cases MLE results in inflated type-I errors, APPS can provide better control.



Supplementary Figure 26: Validation of group-level Type I error rate for group-level test of confirmation bias ($\alpha_c - \alpha_d > 0$) across experiments. The plots show the proportion of false-positive detections (Type I error rate) for a significant group-level positive confirmation bias across ten experimental designs using absolute (**a**) and normalized (**b**) bias metrics. We generated synthetic groups of N agents from a balanced population with zero net generative bias. We then compared the frequency of significant results ($p < 0.05$) across 1,001 iterations using the Anti-Participant Permutation Simulation (APPS) method (green diamonds), direct MLE fitting with a one-tailed t-test (peach squares), and standard MAP fitting with a one-tailed t-test (purple circles). Bars represent the 95% confidence interval. The dashed horizontal line represents the nominal 5% Type I error threshold. Zoomed-in plots around the 5% thresholds are shown in the right. MAP estimation results in systematically inflated false-positive rates across nearly all datasets, reaching near 1 in some experiments (e.g., S1b), whereas both MLE and the APPS method maintain robust Type I error control around the 5% level for these task designs and parameter distributions.

Although our datasets and a majority of existing literature are characterized by positive confirmation biases, for completeness, we sought to explore type-I error rates when detecting negative learning asymmetries (disconfirmation bias: $\alpha_c - \alpha_d < 0$). Thus, we repeated the exact same method but this time calculating 1-sided p-values for the negative side. This revealed that the standard MAP approach produces exceptionally low false-positive rates for disconfirmation bias (Supplementary Fig. 27, violet circles), a result expected given the strong positive estimation biases inherent to that method. Conversely, both MLE (peach squares) and the APPS test (green diamonds) produced Type I error rates were overall near the nominal 5% level for the absolute bias metric (Supplementary Fig. 27a, green diamonds and peach squares; but note a minority of cases where confidence intervals are above the .05). Interestingly, both methods show more consistent inflation of Type I error rates when using the normalized bias metric (Supplementary Fig. 27b, green diamonds and peach squares). Thus, these methods appear biased to detect negative group level normalized effects for null populations.

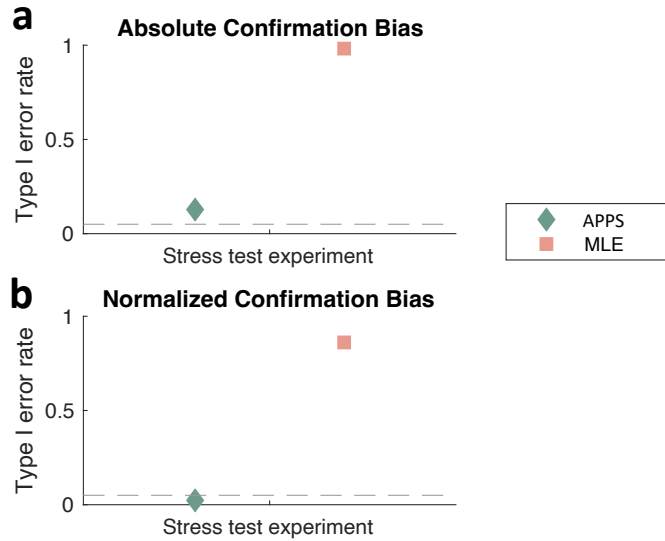


Supplementary Figure 27: Validation of group-level Type I error rate for group-level test of disconfirmation bias ($\alpha_c - \alpha_d < 0$) across experiments The plots show the proportion of false-positive detections (Type I error rate) for a significant group-level negative confirmation bias across ten experimental designs using absolute (a) and normalized (b) bias metrics, generated as the ones in Supplementary Fig. 26. Bars represent the 95% confidence interval. The dashed horizontal line represents the nominal 5% Type I error threshold. Zoomed-in plots around the 5% thresholds are shown in the right. MAP estimation results in systematically minimal false-positive rates across nearly all datasets, and both MLE and the APPS method show Type I error around the 5% level.

3.2.3 Type I error “stress test”

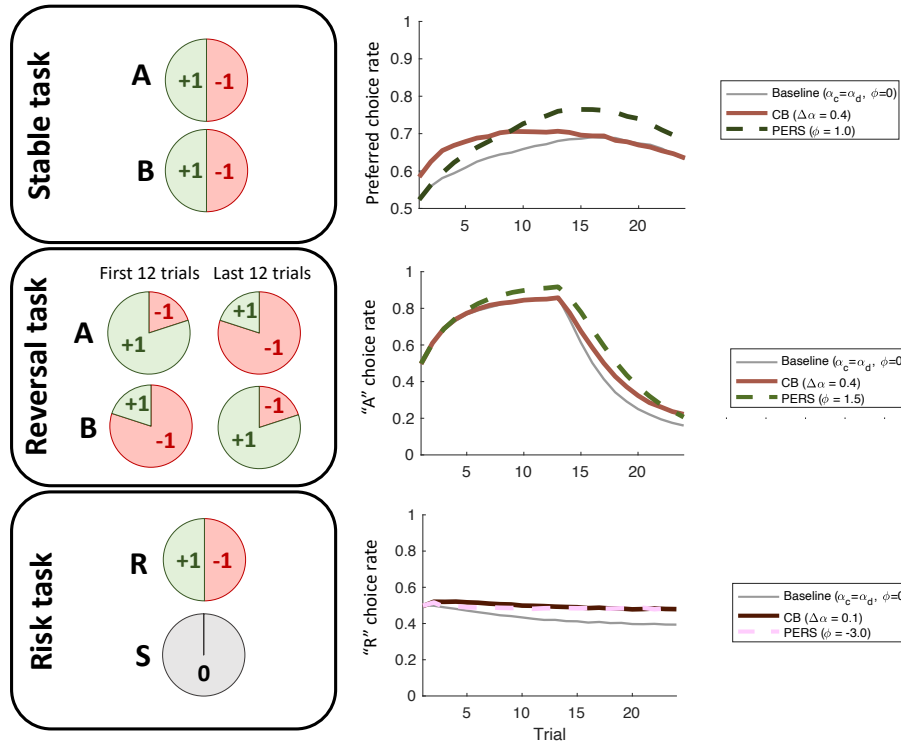
We reasoned that MLE provided an adequate type-I error control for most experiments (Fig S26) because it yielded weak effects in confirmation bias estimates within the parametric ranges represented in our experiments. However, this approach could fail when participants are sampled from ranges where MLE biases are more pronounced. To test this conjecture, we created a synthetic population of 1,001 agents based on a shortened version of the P2 experiment (100 trials instead of the original count) and sampled parameters from ranges previously shown to

maximize model mimicry ($\beta \in U(0,2)$, $\alpha_c \in U(0.3,0.5)$, $\alpha_d \in U(0.3,0.5)$, $\tau \in U(0.4,0.6)$, $\phi \in U(3,5)$; see Supplementary Fig. 21). This specific population had no true net confirmation bias, as the two learning rates were sampled independently. We then repeated the group-level sampling procedure described above, sampling 1001 groups of 200 agents with replacement and testing for a significant positive confirmation bias with either MLE and a t-test, or APPS. This "stress test" revealed that direct MLE fitting was susceptible to fitting biases in these extreme conditions, with false-positive rate of ~100% (or 83% at when measured in normalized terms; Supplementary Fig. 28, peach squares). By contrast, the APPS method provided better control for these biases, producing a notably lower Type I error rate (12.7% in absolute terms, and 2.5% in normalized terms; Supplementary Fig. 28, green diamonds). While our method significantly outperformed standard testing procedures, the fact that the error rate was above 5% (when measured in absolute terms) in this extreme scenario suggests researchers should always evaluate the statistical power and validity of their specific methodologies, within the context of their unique task design and population sample.

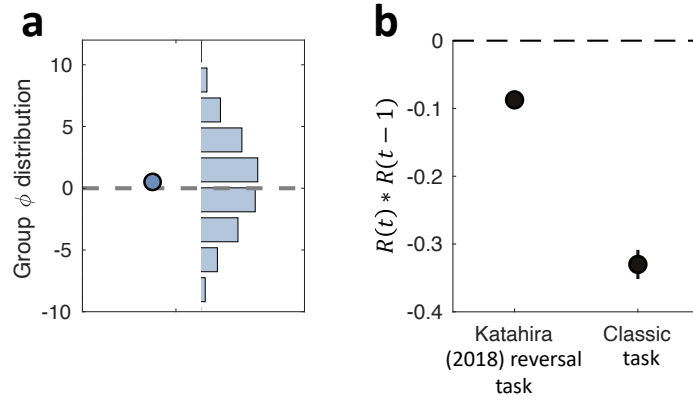


Supplementary Figure 28: Group-level Type I error rates in a high-mimicry stress test. Comparison of Type I error rates for direct MLE fitting (peach squares) and the APPS method (green diamonds) in a shortened experiment (100 trials) using a synthetic population with parameters optimized to maximize model mimicry and fitting biases. **(a)** Absolute confirmation bias metrics. Under these extreme conditions, MLE fails completely with a 100% false-positive rate, while the APPS method maintains much greater robustness (12.7%). **(b)** Normalized confirmation bias metrics. The APPS method effectively controls for the artifact, yielding a Type I error rate of 2.5%, compared to 83% for direct MLE. The dashed line represents the 5% alpha threshold. Error bars (mostly covered by the main markers) represent 95% confidence intervals. Normalized confirmation bias metrics. The APPS method effectively controls for the artifact, yielding a Type I error rate of 2.5%, compared to 83% for direct MLE. The dashed line represents the 5% alpha threshold. Error bars (mostly covered by the main markers) represent 95% confidence intervals.

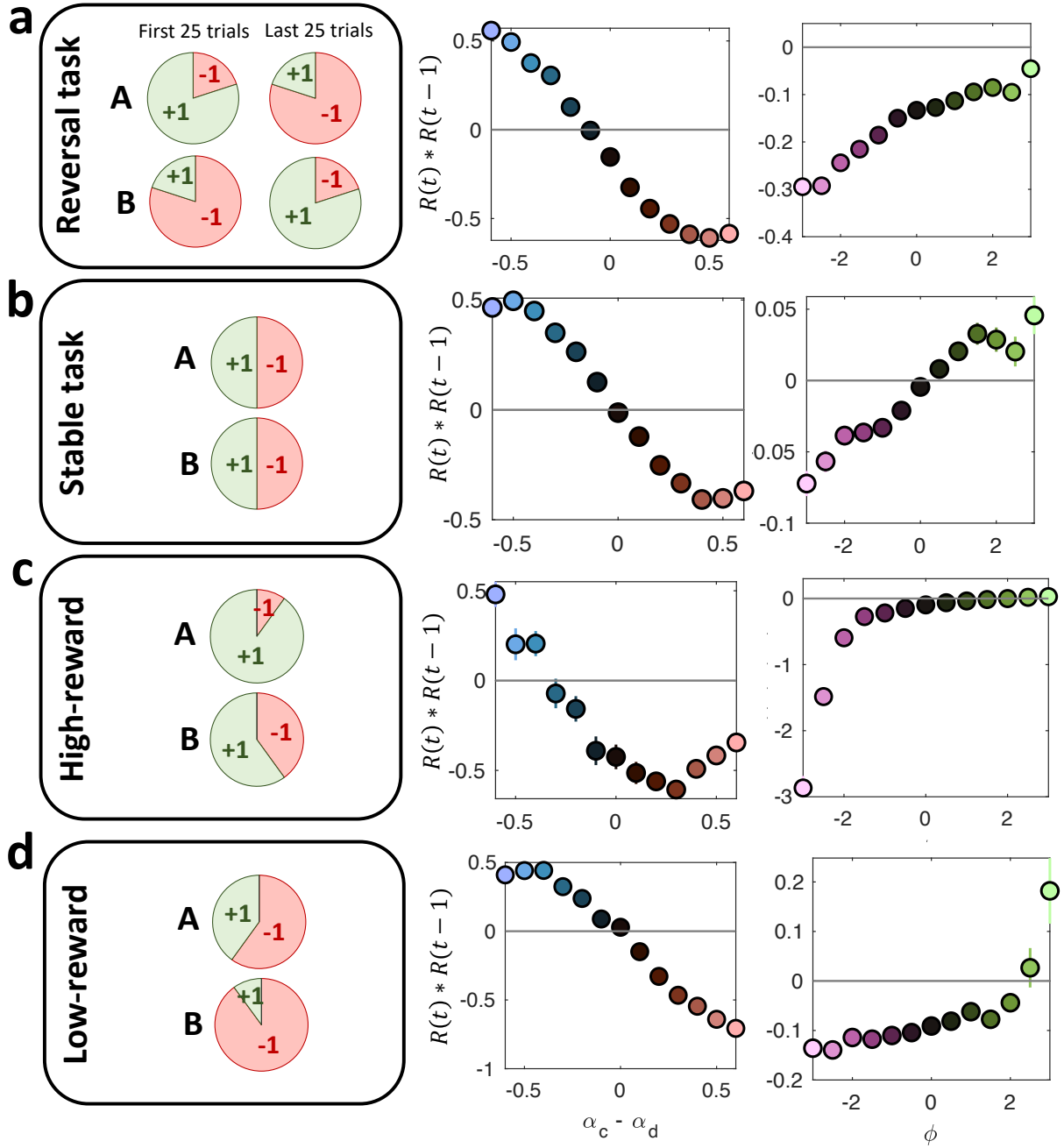
Supplementary note 4: Behavioral signature results



Supplementary Figure 29: Direct comparison of levels of confirmation bias and perseveration producing similar behavioural signatures. The analysis is performed in the same way as the one in the main text, but plotting the choice rate lines on top of each other to facilitate comparability. Thin lines represent the choice rate of a baseline agent with symmetric learning and without perseveration. Solid thick lines represent an agent with confirmation bias (of the magnitude specified in legend), but without perseveration. Dashed thick lines represent a PSL agent (with the perseveration magnitude specified in legend). **(a)** Stable Task. Both LA and PSL agents produce a curve above the baseline. **(b)** Reversal Task. Both LA and PSL agents predict a slower adaptation to reversal than baseline. **(c)** Risk Task. Both LA and PSL agents predict higher risk-seeking behaviour than baseline.



Supplementary Figure 30: Robustness of Katahira (2018) signature for a population with symmetric learning and average positive perseveration. Our previous parameter sweeps (Fig 6b top-right) demonstrated that while positive and negative perseveration tend to produce positive and negative regression interaction term, respectively, the former effects are stronger ($\phi < 0$). We reasoned as a consequence of this negative effect domination, a population with symmetric learning and an overall positive perseverative tendency can still produce a spurious negative interaction term (the putative signature of confirmation bias). **(a)** Generative parameter distribution (dot represents mean perseveration). We simulated 1,000 groups of 20 participants. For each participant, parameters were fixed at the average PSL best-fit estimates from experiment P2 ($\beta = 1.79, \alpha = 0.49, \tau = 0.38$), with the exception of the perseveration weight (ϕ), which was sampled from a normal distribution with a positive mean ($\phi \sim N(0.5, 4)$). **(b, left)** Reversal task (Katahira, 2018) interaction terms. Average of the $R(t) * R(t-1)$ interaction terms recovered from Katahira’s (2018) reversal task, across the 1,000 groups, using a mixed-effects logistic regression model with random intercepts ($stay \sim R(t) * R(t-1) + (1|participant)$). **(b, right)** Same as b (right) but for a classic (static, non-reversal) bandit task. In both experimental designs, the interaction terms were significantly negative, demonstrating that even in populations positive perseveration, regression-based signatures can incorrectly signal the presence of a confirmation bias. Dots and error bars represent the mean and s.e.m. of the interaction terms across the 1000 synthetic datasets.



Supplementary Figure 31: 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.

SUPPLEMENTARY TABLES

A Hybrid Model Reduces, but Does Not Eliminate, Measured Confirmation Bias in Empirical Data

Supplementary Table 1: Group-level confirmation bias in participants based on Pure Learning Asymmetry (LA) model fits with MAP. One-sided t -tests ($\alpha_c - \alpha_d > 0$) show a significant confirmation bias across all ten experiments under MAP estimation, prior to controlling for choice perseveration.

Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	5.605	<0.001	0.793	[0.124, 0.262]
L2	34	5.435	<0.001	0.919	[0.187, 0.411]
P1	19	2.364	0.029	0.529	[0.020, 0.326]
P2	19	11.713	<0.001	2.619	[0.309, 0.443]
C1	23	2.611	0.016	0.533	[0.036, 0.310]
C2	23	9.753	<0.001	1.991	[0.262, 0.403]
C3	29	3.155	0.004	0.576	[0.063, 0.295]
C4	19	4.926	<0.001	1.102	[0.108, 0.269]
S1a	142	4.815	<0.001	0.403	[0.104, 0.248]
S1b	142	12.061	<0.001	1.009	[0.236, 0.329]

Supplementary Table 2: Group-level confirmation bias in participants based on the Hybrid model fits with MAP. One-sample t -tests vs zero ($\alpha_c - \alpha_d > 0$) evaluate the strength of the estimated learning asymmetry after accounting for choice perseveration. A statistically significant confirmation bias ($p < 0.05$) survives in six out of the ten bandit-task experiments (L1, L2, P2, C2, C4, and S1b).

Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	2.336	0.024	0.330	[0.014, 0.189]
L2	34	2.396	0.022	0.405	[0.027, 0.325]
P1	19	0.869	0.396	0.194	[-0.084, 0.204]
P2	19	7.616	<0.001	1.703	[0.239, 0.421]
C1	23	1.827	0.081	0.373	[-0.018, 0.282]
C2	23	9.003	<0.001	1.838	[0.212, 0.339]
C3	29	1.055	0.300	0.193	[-0.071, 0.223]
C4	19	3.522	0.002	0.788	[0.083, 0.325]
S1a	142	0.261	0.794	0.022	[-0.070, 0.091]
S1b	142	4.963	<0.001	0.415	[0.067, 0.156]

Supplementary Table 3: Comparison of estimated confirmation bias between the Pure Learning Asymmetry (LA) and Hybrid model MAP fits of participants behavior. Paired t -tests directly compare the learning rate asymmetry estimates ($\alpha_c - \alpha_d$) obtained from the LA model versus those from the hybrid model. Accounting for choice perseveration in the hybrid model leads to a statistically significant reduction ($p < 0.05$) in measured confirmation bias across six out of ten experiments. For

the remaining four experiments (P2, C1, C3, and C4), the reduction is not statistically significant. This highlights how omitting perseveration parameters systematically overestimates the true extent of confirmation bias.

Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	2.499	0.016	0.353	[0.018, 0.165]
L2	34	2.582	0.014	0.436	[0.026, 0.221]
P1	19	3.418	0.003	0.764	[0.044, 0.182]
P2	19	1.239	0.230	0.277	[-0.032, 0.123]
C1	23	0.864	0.396	0.176	[-0.056, 0.137]
C2	23	3.300	0.003	0.674	[0.021, 0.092]
C3	29	1.377	0.179	0.251	[-0.050, 0.256]
C4	19	-0.330	0.745	-0.074	[-0.110, 0.080]
S1a	142	6.585	<0.001	0.551	[0.116, 0.215]
S1b	142	7.480	<0.001	0.626	[0.126, 0.216]

Choice perseveration masquerades as spurious confirmation bias

Supplementary Table 4: Spurious confirmation bias recovered from fitting perseverative, symmetrically learning agent simulations with the hybrid model and MAP. One-sample t-tests vs zero ($\alpha_c - \alpha_d > 0$) evaluate the spurious learning rate asymmetry detected when fitting simulated datasets from the Perseverative Symmetrical Learning (PSL) model with the hybrid model. Because the generative agents learn symmetrically, any significant deviation from zero represents an artifactual confirmation bias. Under standard MAP fitting conditions, a highly significant spurious confirmation bias ($p < 0.05$) is detected in nine out of ten experiments (all except C4), illustrating the severe impact of model mimicry.

Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	6.174	<0.001	0.873	[0.039, 0.076]
L2	34	6.338	<0.001	1.071	[0.049, 0.096]
P1	19	4.200	<0.001	0.939	[0.029, 0.086]
P2	19	7.703	<0.001	1.723	[0.115, 0.200]
C1	23	3.794	0.001	0.774	[0.011, 0.037]
C2	23	5.583	<0.001	1.140	[0.056, 0.121]
C3	29	3.842	0.001	0.701	[0.022, 0.071]
C4	19	1.907	0.072	0.426	[-0.002, 0.035]
S1a	142	6.196	<0.001	0.518	[0.019, 0.038]
S1b	142	15.169	<0.001	1.269	[0.095, 0.124]

Supplementary Table 5: Underestimation of the choice perseveration parameter (ϕ) in the hybrid model compared to the PSL model based on empirical fits with MAP. Paired t-tests directly compare the estimated choice perseveration weight (ϕ) obtained from the hybrid model versus the perseverative symmetrical learning (PSL) model. Except for experiment C4, the estimated perseveration parameter is

significantly smaller in the hybrid model across all datasets. This systematic underestimation indicates a parameter-estimation trade-off, where the hybrid model absorbs some of the variance from outcome-independent perseveration into spurious confirmation bias.

Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	-5.585	<0.001	-0.790	[-0.742, -0.349]
L2	34	-4.949	<0.001	-0.836	[-0.731, -0.305]
P1	19	-3.346	0.003	-0.748	[-0.681, -0.157]
P2	19	-6.551	<0.001	-1.465	[-1.759, -0.907]
C1	23	-4.172	<0.001	-0.852	[-1.136, -0.383]
C2	23	-9.457	<0.001	-1.930	[-2.592, -1.661]
C3	29	-2.764	0.010	-0.505	[-0.649, -0.097]
C4	19	-0.269	0.791	-0.060	[-0.208, 0.160]
S1a	142	-4.726	<0.001	-0.395	[-0.274, -0.112]
S1b	142	-8.568	<0.001	-0.716	[-0.572, -0.357]

Supplementary Table 6: Recovered choice perseveration parameter (ϕ) when fitting simulated PSL agent data with the hybrid model using MAP. One-sample t-tests vs zero show that a highly significant, positive perseveration weight ($\phi > 0$) is recovered across all ten experiments. While the parameter is robustly detected, the recovered values are systematically underestimated compared to their true generative parameters due to the variance-partitioning trade-offs inherent to the hybrid fitting architecture.

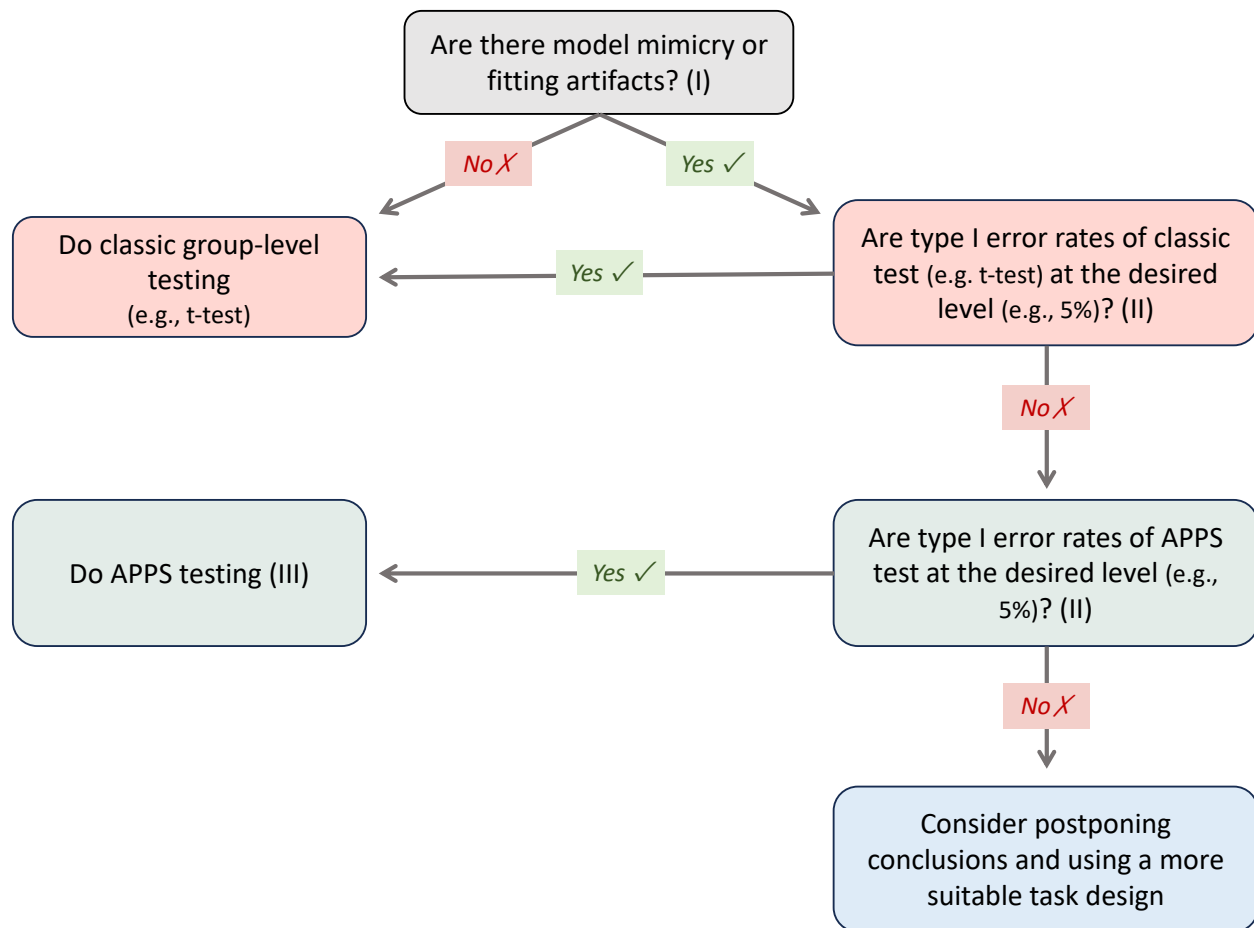
Experiment	df	t-value	p-value	Cohen's d	95% CI
L1	49	8.313	<0.001	1.176	[0.428, 0.700]
L2	34	9.739	<0.001	1.646	[0.565, 0.862]
P1	19	6.906	<0.001	1.544	[0.579, 1.082]
P2	19	9.823	<0.001	2.196	[0.609, 0.938]
C1	23	7.535	<0.001	1.538	[0.875, 1.537]
C2	23	11.816	<0.001	2.412	[0.867, 1.235]
C3	29	5.358	<0.001	0.978	[0.440, 0.983]
C4	19	9.810	<0.001	2.194	[0.923, 1.423]
S1a	142	13.100	<0.001	1.096	[0.500, 0.677]
S1b	142	17.543	<0.001	1.467	[0.549, 0.689]

SUPPLEMENTARY METHODS

The Modeler's Handbook: A Guide to Addressing Mimicry Problems

The central challenge in many areas of computational modeling is distinguishing between two mechanisms, A and B, that produce similar behavioral outputs. For instance, in the context of our study, this manifests as a confusion between learning asymmetries (mechanism A) and choice perseveration (mechanism B), both of which can lead to an increase/decrease choice repetition. When a researcher wishes to test for the significance of mechanism A while controlling for the potential influence of mechanism B, they must account for "partial model mimicry." This occurs when parameter estimates regarding the focal mechanism A are biased due to contributions from mechanism B. This, in turn, may lead to inflated type-I error rates for standard statistical tests conducted on process-A parameter estimates. This handbook outlines a recommended protocol for diagnosing these problems and applying a suitable statistical test approach to mitigate this problem, taking into account that standard approaches are usually simpler to use (e.g., standard t-test compared to APPS).

As illustrated in the decision flow (Fig.S32), the process begins with a diagnosis (Method I) to determine if the task design and model architecture are susceptible to fitting artifacts or parameter leakage. If no mimicry is detected, classic group-level testing, such as a t-test on estimated parameters, is sufficient. However, if mimicry is present, one must validate the Type I error rate of the chosen statistical test (Method II applied to a simple t-test). If the classic test shows an adequate type I error rate, one can adopt it as the primary inferential tool. Conversely, if the classic test shows an inflated Type I error rate (e.g., exceeding the nominal 5% level), the same validation procedure is applied to the Anti-Participant Permutation (APPS) test (Method II applied to APPS). If the APPS test successfully recovers the intended alpha level, it is adopted as the primary inferential tool (Method III); if both methods fail to control for false positives, we recommend switching to a more suitable task design. Of course, researchers should always aim to use task designs that provide adequate control for potential confounds. However, often researchers become aware of potential confounds only after data-collection and may wish to still statistically control for these.



Supplementary Figure 32: Hierarchical decision pipeline for parameter-based group-level inference while controlling for confounds due to model mimicry and fitting artifacts. The process begins with method I (diagnosis), assessing if the task and model combination are susceptible to parameter leakage. If mimicry is absent, standard group-level tests (e.g., t-test on MLE parameters) are appropriate. If mimicry is detected, the process moves to method II (validation) to evaluate the Type I error rates of classic tests. If these rates are inflated (e.g., exceeding 5%), the validation is repeated for the Anti-Participant Permutation Simulation (APPS) test. Successful validation leads to Step III (APPS testing) for final inference. If neither method properly controls for false positives, the researcher is encouraged to reconsider alternative design that offer better separability between the processes.

Method I. Is there a model mimicry/fitting bias problem?

The first step in any modeling pipeline should be to determine whether the chosen model and task design are susceptible to spurious detections. One approach for testing that is to generate simulations of the nuisance mechanism (B) alone and test whether these simulated data “show traces” of the focal process A. To do this, follow these steps:

1. **Fit the empirical behavioral data using a restricted model that only includes the nuisance mechanism (B).**
2. **Use these parameters to simulate a large number of synthetic datasets based on the model that only includes mechanism B.** To ensure high sensitivity, we recommend generating multiple simulations (for instance, $n=10$) for every individual participant's

parameter set. These synthetic agents represent a version of the original population under the null hypothesis that they only present the confound (i.e., that they present as much of the confound as possible).

3. **Re-fit these synthetic dataset using a "hybrid" model that includes both the focal mechanism (A) and the nuisance mechanism (B).** Because the generative process strictly lacked mechanism A, any significant effect recovered for parameter A is, by definition, a model-fitting artifact or "parameter leakage."

TEST CONCLUSION: If the hybrid model recovers significant results for the focal parameter from these confound-only simulations, the conclusion is that ***"model mimicry and fitting artifacts might be producing spurious results for the focal parameter of interest in this specific dataset"***.

This exercise is essential because the degree of between process leakage might be highly sensitive to the trial count of the task, the specific task design, and the distribution of parameters in the target population. If a model-mimicry problem does exist, one can try alternative fitting procedures to test if mimicry is eliminated (e.g., what happens if you change or even remove parameter priors?).

Method II. Type I Error Validation. Does a test lead to an inflated false positive rate?

At several stages in the pipeline we need to validate that the chosen group-level test, whether it be the APPS test (method III) or a simpler alternative, is performing correctly for the given research context, or whether it has an inflated false positive rate. We can approximate such type I errors in the following way:

1. **Construct a Balanced Pseudo-Population:** Create a "super-population" of $2N$ pseudo-agents by duplicating your original participant pool (N , based on hybrid model fits) and adding their mirrored anti-participants.
 - **The Pseudo-agents:** Uses the recovered parameters from original participants as-is to simulate the hybrid model.
 - **The Pseudo-anti-agents:** Uses identical parameters, but the focal mechanism that we want to test is "mirrored" around 0 (e.g., swapping α_c and α_d if our null hypothesis is $\alpha_c - \alpha_d = 0$, changing ϕ for $-\phi$ if null hypothesis is $\phi = 0$).

Note: in our examples we want to test the mechanism against a null of 0, but this test can be done to test against any other value, by just mirroring around that value. For instance, if we want to test for a ϕ value of 0.5, we can create anti-agents by replacing the perseveration parameter with $0.5 - (\phi - 0.5)$.

By design, this pseudo-population has a net focal effect matching exactly the null, ensuring that any significant result for samples from this pool is a false positive. This population preserves the empirical distribution of non-focal parameters (including those related to the confounding mechanism), maintaining realistic confounding effects.

2. **The Pseudo-Validation Sampling Loop:** To estimate the global error rate, repeat the following procedure many times (e.g., 1,000 iterations):

- **Sample Pseudo-Participants:** Draw N pseudo-agents (with replacement) from the 2N pseudo-population. These are your "pseudo-participants".
- **Generate Pseudo-Empirical Effects:** For each pseudo-participant, forward-simulate a pseudo-behavioral dataset of the hybrid model using the experimental task and fit it with the hybrid model. Average these estimates across the N pseudo-participants to obtain a simulated "pseudo-empirical" group effect under the null hypothesis.

Computational Efficiency Note: Performing these simulations de novo for every sampling iteration is computationally intensive. To optimize this, researchers could pre-generate a pool of simulated pseudo-participants (e.g., 30 per pseudo-agent; 2Nx30) and their corresponding model fits. When a sample is drawn, we randomly selected one of these pre-pseudo-datasets for each pseudo-participant to calculate the pseudo-empirical group-level effect size. If one wants to test a simple frequentist approach (e.g., t-test) one can apply it directly to these pseudo-empirical effects of the sampled participants.

- **Apply the Pseudo-Statistical Test:** Run the full statistical procedure (e.g., the APPS test, see step III) on these pseudo-data as if they were real empirical data. This involves generating a null distribution, permuting between model-fits to simulated behavior for their pseudo-agent and pseudo-anti-agent simulations (the "pseudo-null-effects"), and calculating a pseudo-p-value for the simulated pseudo-empirical group effect.

Computational Efficiency Note: As in the previous point, performing these simulations de novo for every sampling iteration is computationally intensive. To optimize this, researchers could pre-calculated the corresponding null effects for the APPS tests by simulating and fitting pseudo-agents and pseudo-anti-agents for each pseudo-dataset (e.g., 30 per pseudo-dataset; 2Nx30x2x30 pseudo-null-effects). When computing the APPS test on a sample, one should simply permuted over these pre-calculated values for the N sampled pseudo-participants (each with 2x30 precomputed pseudo-(anti)agent null values) to generate the null distribution.

- **Record Rejections:** Record whether the test rejected the null hypothesis (e.g., pseudo-p < 0.05).

3. **Calibrate the Alpha Level:** The final Type I error rate is the proportion of iterations in which the null was incorrectly rejected. If the rate of these pseudo-rejections is significantly higher than the intended 5%, the method doesn't provide adequate type-I error control.

TEST CONCLUSION: "*Under the assumption that the synthetic population is a good approximation of a null effect, we can conclude that the chosen statistical test produces a specific Type I error rate of x%.*"

Method III. The Anti-Participant Permutation Simulations test (APPS test)

If the classic statistical test produces inflated type I error rates, one must create a custom null distribution that "pre-adjusts" for model mimicry and fitting biases. One way to approximate such a distribution is using the APPS test proposed in this study:

1. **Obtain the empirical group mean:** Fit your actual participants with the hybrid model (MLE) to get the empirical group-level mean of the learning asymmetry.

2. **Create the "Agent + Anti-Agent" Population:** For every participant, create two artificial agents based on their hybrid fits:
 - **The Agent:** Uses the recovered parameters as-is.
 - **The Anti-Agent:** Uses identical parameters, but the focal mechanism that we want to test is "mirrored" around 0 (e.g., swapping α_c and α_d if our null hypothesis is $\alpha_c - \alpha_d = 0$, changing ϕ for $-\phi$ if null hypothesis is $\phi = 0$). This ensures that the *average* generative parameters for the "population of agents and anti agents" equals the value stated in the null hypothesis we want to test.

Note: in our examples we want to test the mechanism against a null of 0, but this test can be done to test against any other value, by just mirroring around that value. For instance, if we want to test for a ϕ value of 0.5, we can create anti-agents by replacing the perseveration parameter with $0.5 - (\phi - 0.5)$.
3. **Map the Artifact Field:** Generate N simulated datasets of the hybrid model for every agent and every anti-agent (e.g., N=1001). Re-fit all of them with the hybrid model.
4. **The Two-Stage Permutation Sampling:** Perform M iterations to build the group null distribution (we recommend M=10000). In each iteration:
 - **First**, for each of the N participants, randomly select *either* the agent *or* the anti-agent.
 - **Second**, randomly draw one of that selected member's 1,001 re-fitted estimates.
 - Calculate the mean of these N selections.
5. **Statistical Inference:** Your p-value is the proportion of these effect means (from step 4) that are greater than or equal to your empirical group mean (from step 1). This generates a p-value that is, in principle, already "pre-adjusted" for the specific fitting artifacts and model mimicry issues of the study.

IMPORTANT: this exercise should be performed for every experiment (i.e., population sample performing a specific task), since the fitting biases and model leakage can change depending on the task design and parameter distribution. One should not re-use null distributions for a new task or population sample.

TEST CONCLUSION: If a significant result is obtained, we conclude that "***the observed empirical effect is significantly greater than the spurious bias expected from model mimicry and fitting artifacts alone, as estimated by the APPS procedure***".