Decomposing Loss Aversion: Valuation vs. Attention in a Decision Field Theory Framework

Akhil Abburu (akhil.abburu@hss.iitd.ac.in)[1] **& Sumeet Agarwal**[2]
[1]Department of Humanities and Social Sciences, Indian Institute of Technology Delhi
[2]Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi
New Delhi, India

Abstract

Loss aversion—the idea that losses loom larger than equivalent gains—is a cornerstone of behavioral economics. Yet estimates of the loss aversion parameter (λ) vary considerably across elicitation methods, suggesting this single parameter may conflate distinct psychological mechanisms. We test this hypothesis using a hybrid Decision Field Theory, which allows formal separation of valuation asymmetry (λ: losses weighted more heavily in utility) from attentional asymmetry (κ: losses sampled more frequently during deliberation). Fitting hierarchical Bayesian models to 10 datasets (N = 686; ~ 140,000 trials), we find that attentional asymmetry alone provides superior predictive accuracy in the majority of cases. Simulation analyses further demonstrate that standard utility models absorb attentional variance into inflated λ estimates. These findings suggest that the tendency to reject favorable mixed gambles may substantially reflect how losses are weighted during accumulation, rather than solely due to asymmetric valuation.

Keywords: loss aversion; computational modeling; decision field theory; sequential sampling models

The tendency to reject favorable mixed gambles, where the potential loss is smaller than the potential gain, is a robust phenomenon in decision-making (Ruggeri et al., 2020). Since the development of Prospect Theory (PT; Kahneman and Tversky, 1979), the dominant explanation for this behavior has been *loss aversion* (λ): a stable, intrinsic valuation asymmetry where "losses loom larger than gains". Under this view, the asymmetry resides in the psychophysics of the utility function itself, a "kink" at the reference point that renders negative outcomes subjectively steeper than equivalent positive ones.

However, recent meta-analyses and large-scale replications have challenged the stability of λ as a unitary trait (Gal and Rucker, 2018; Walasek et al., 2021; Yechiam and Hochman, 2013; see also Spektor et al., 2024). Estimates of λ fluctuate significantly depending on the elicitation method, the range of outcomes, the magnitude of the outcomes, and the estimation method (Abburu et al., 2025; Mukherjee et al., 2025; Walasek and Stewart, 2015; see also Brown et al., 2023; Mrkva et al., 2020). Additionally, process-tracing studies suggest that this instability may reflect a dynamic allocation of resources rather than a static preference, where loss aversion arises from a selective attentional bias toward negative outcomes rather than a stable valuation asymmetry (Lejarraga et al., 2019; Pachur et al., 2018). This instability raises a fundamental question: if λ cannot be reliably estimated, what explains the robust behavioral phenomenon it was meant to capture? One possibility is that static utility models conflate distinct psychological mechanisms into a single parameter.

Sequential sampling frameworks offer a path toward mechanistic decomposition. Zhao et al. (2020) applied Drift Diffusion Modeling (DDM) to mixed gambles, finding that rejection reflects not only asymmetric valuation but also a pre-valuation bias favoring rejection. Similarly, Clay et al. (2017) localized loss aversion in drift rate (information uptake) rather than response bias. Sheng et al. (2020) extended this approach with eye-tracking, showing that preferential gaze toward losses indexes a distinct attentional component. These DDM-based decompositions illustrate that unitary "loss aversion" may reflect multiple cognitive sources, but they rely on response times or eye-tracking data.

Decision Field Theory (DFT; Busemeyer and Townsend 1993) provides an alternative decomposition. Unlike DDM, DFT derives diffusion variance endogenously from conflict between option values, allowing valuation (λ) and attentional (κ) asymmetries to leave distinguishable signatures in choice probabilities themselves. Although sequential sampling models are most often applied to response-time data, an established literature has used DFT to adjudicate among accounts of preferential choice from choices alone—probabilistic DFT outperforming CPT and the priority heuristic on binary risky choices (Rieskamp, 2008), a minor DFT extension capturing canonical CPT paradoxes (Bhatia, 2014), and DFT$_\epsilon$ providing the most parsimonious account when CPT, TAX, and DFT$_\epsilon$ are compared head-to-head (Kellen et al., 2020). We extend this tradition by asking whether DFT's choice-only signatures can adjudicate the two mechanisms that have been advanced for loss aversion.

We address this gap by integrating a Prospect Theory-style valuation parameter (λ) directly into the drift rate structure of a Hierarchical Bayesian DFT. This allows us to test two competing mechanisms for mixed-gamble rejection within a single generative framework: (1) Valuation Asymmetry, where the objective magnitude of a loss is transformed into a larger subjective utility (λ); versus (2) Attentional Asymmetry, where the decision-maker allocates disproportionate attentional weight to loss outcomes (κ) during evidence accumulation. We fit four nested model variants—Null (M_0), Valuation (M_λ), Attention (M_κ), and a Full Model ($M_{\lambda\kappa}$)—to 10 datasets (N = 686, approximately 140,000 trials). We restrict inference to choice data alone, exploiting DFT's endogenous

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

variance structure to render these mechanisms identifiable; parameter recovery shows that model selection correctly distinguishes κ- from λ-generated data in both directions (see Results).

Theoretical and Computational Framework

We employ a Decision Field Theory (DFT) framework to decompose the cognitive mechanisms of loss aversion. DFT conceptualizes deliberation not as a static calculation of Expected Utility, but as a sequential accumulation of evidence. Faced with a gamble, the decision-maker samples affective evaluations ("valences") of potential outcomes over time. These evaluations accumulate into a preference state $P(t)$, drifting toward an acceptance or rejection threshold.

To ensure computational tractability for hierarchical Bayesian estimation, we utilize the analytical solution of the Wiener diffusion process. In this framework, the probability of accepting a gamble is a function of the signal-to-noise ratio of the accumulation process. Specifically, the log-odds of acceptance are given by:

$$\ln\left(\frac{P(\text{accept})}{P(\text{reject})}\right) = \frac{2 \cdot \epsilon \cdot \delta}{\sigma^2} \tag{1}$$

where δ is the drift rate (the expected valence), σ^2 is the diffusion variance, and ϵ is a scaling parameter representing the decision threshold (analogous to boundary separation a in standard drift-diffusion models). Crucially, unlike standard models where drift and variance are independent free parameters, DFT defines them *endogenously* based on the subjective utilities and attention weights of the gamble's outcomes. This constraint provides a potential basis for distinguishing valuation from attention effects.

Utility and Valuation

We assume a subjective utility function $u(x)$ governed by a curvature parameter γ and a valuation weighting λ. Following Cumulative Prospect Theory, we define the utility of the gain (g) and loss (l) magnitudes as:

$$u(g) = g^{\gamma}, \quad u(l) = -\lambda \cdot l^{\gamma} \tag{2}$$

Here, λ captures *Valuation Asymmetry*. If $\lambda > 1$, the negative utility generated by a loss is magnified relative to an equivalent gain. This corresponds to the traditional definition where losses feel larger due to the steepness of the utility function.

Attention and Drift

The drift rate δ is the expectation of the momentary valences. It is determined by the weighted sum of utilities, where the weights reflect the allocation of attention:

$$\delta = W_g \cdot u(g) + W_l \cdot u(l) \tag{3}$$

We parameterize attention via a bias parameter $\kappa \in (-1, 1)$. The attention weight allocated to the loss outcome is:

$$W_l = 0.5 + 0.5\kappa \tag{4}$$

and $W_g = 1 - W_l$. Here, κ captures *Attentional Asymmetry*. When $\kappa = 0$, attention is symmetric ($W_l = W_g = 0.5$). When $\kappa > 0$, the decision-maker systematically over-samples the loss outcome ($W_l > 0.5$), driving the drift rate toward rejection regardless of utility curvature.

κ differs from the distraction parameter in prior DDFT work (Bhatia, 2014). Distraction adds uniform noise across events; κ directionally re-weights channels. When $\kappa > 0$, losses receive more weight in drift computation, increasing their contribution to evidence accumulation without requiring value function distortion.

Endogenous Variance and Identifiability

A key theoretical challenge is the conflation of these mechanisms: a shift in λ and a shift in κ can both lower the drift rate δ, reducing the probability of acceptance. However, DFT creates a structural constraint that renders them identifiable: the diffusion variance σ^2 is defined as the variance of the weighted valence differences:

$$\sigma^2 = W_g \cdot W_l \cdot (u(g) - u(l))^2 \tag{5}$$

This variance term appears in the denominator of the choice rule (Eq. 1), creating distinct signatures for each mechanism. Increasing λ amplifies the utility difference $|u(g) - u(l)|$, thereby increasing σ^2 and reducing choice consistency. Conversely, shifting κ away from zero moves attention away from the symmetric allocation that maximizes $W_g \cdot W_l$, thereby decreasing σ^2 and increasing choice consistency. By estimating λ and κ within this framework, we can assess whether rejection behavior is better explained by valuation versus attention.

Methods

Datasets

We analyzed 10 datasets containing equiprobable mixed gambles from seven published studies (Khan et al., 2024; Sheng et al., 2020; Sokol-Hessner, Hartley, et al., 2015; Sokol-Hessner, Lackovic, et al., 2015; Sokol-Hessner et al., 2009, 2016; Zhao et al., 2020), comprising $N = 686$ participants and approximately 140,000 binary choice trials. Across all trials, participants decided whether to accept or reject gambles offering a 50% chance of winning (losing) a specified gain (loss).

Data Preprocessing

Trials were excluded if they contained missing responses or invalid stimulus values. At the participant level, we excluded individuals with fewer than 20 valid trials or extreme response patterns (acceptance rates below 5% or above 95%), as such patterns preclude meaningful estimation of individual-level parameters. This yielded a sample suitable for hierarchical estimation while ensuring sufficient within-subject variance for parameter identification.

Computational Model and Variants

We implemented the DFT framework (Equations 1–5) as a hierarchical Bayesian model in Stan. The choice probability

follows directly from Equation 1; we describe here only the hierarchical structure and estimation procedures.

We estimated four nested model variants to isolate the contributions of valuation and attentional asymmetries (Table 1). The Null Model (M_0) assumes symmetric valuation ($\lambda = 1$) and symmetric attention ($\kappa = 0$), attributing all individual differences to the scaling parameter ϵ. The Valuation Model (M_λ) freely estimates λ while fixing $\kappa = 0$, testing whether loss aversion arises from asymmetric utility. The Attention Model (M_κ) fixes $\lambda = 1$ and estimates κ, testing whether loss aversion reflects biased information sampling. The Full Model ($M_{\lambda\kappa}$) estimates both parameters simultaneously to assess whether both mechanisms contribute independently. Across all models, utility curvature was fixed at $\gamma = 0.88$, consistent with canonical Prospect Theory estimates (Tversky & Kahneman, 1992).

Table 1: Model Variants and Free Parameters

Model	ϵ	λ	κ	Interpretation
M_0	Free	Fixed	Fixed	No asymmetry
M_λ	Free	Free	Fixed	Valuation bias
M_κ	Free	Fixed	Free	Attention bias
$M_{\lambda\kappa}$	Free	Free	Free	Both mechanisms

Hierarchical Structure

All models employed a hierarchical (multilevel) structure to partially pool information across participants. Each parameter was decomposed into a group-level mean, group-level standard deviation, and individual-level deviations. This approach balances the bias-variance tradeoff: participants with limited data are regularized toward the group mean, while participants with robust data are allowed to deviate. To avoid the 'funnel' pathologies common in hierarchical cognitive modeling, we used non-centered parameterization for all group-level variance components to facilitate efficient sampling (Papaspiliopoulos et al., 2007).

Parameters were constrained to theoretically meaningful ranges via link functions: $\epsilon_i = \exp(\mu_{\log\epsilon} + \sigma_{\log\epsilon} \cdot z_i^{(\epsilon)})$ (log-normal, ensuring $\epsilon > 0$), $\lambda_i = 0.01 + 4.99 \cdot \text{logit}^{-1}(\mu_\lambda^* + \sigma_\lambda^* \cdot z_i^{(\lambda)})$ (scaled logit-normal, $\lambda \in [0.01, 5.0]$), and $\kappa_i = 0.99 \cdot \tanh(\mu_\kappa^* + \sigma_\kappa^* \cdot z_i^{(\kappa)})$ (scaled tanh-normal, $\kappa \in (-0.99, 0.99)$).

Weakly informative priors were placed on group-level parameters. The prior on λ corresponds to a prior median of approximately $\lambda = 1.5$, encompassing both loss-neutral and moderately loss-averse values. The prior on κ is centered at zero (symmetric attention) with sufficient dispersion to detect biases in either direction.

Estimation

Models were implemented in the probabilistic programming language Stan Carpenter et al., 2017 via the `cmdstanr` interface (Gabry et al., 2025). We ran 4 parallel chains with 3,000 warmup iterations and 5,000 sampling iterations per chain (yielding 20,000 posterior draws per model). Sampler parameters were tuned per model: M_0 used `adapt_delta` = 0.90; M_λ and M_κ used `adapt_delta` = 0.95 with `max_treedepth` = 11; $M_{\lambda\kappa}$ used `adapt_delta` = 0.95 with `max_treedepth` = 12. Convergence was assessed via the $\hat{R}$ statistic (all $\hat{R} < 1.01$), ESS, and through traceplots.

Model Comparison

We compared models using approximate leave-one-out cross-validation (LOO-CV) computed via Pareto-smoothed importance sampling (PSIS-LOO; Vehtari et al., 2017). When Pareto-k diagnostics exceeded 0.7 for individual observations (indicating unreliable importance weights), we applied moment matching to stabilize estimates. Model selection was based on the expected log pointwise predictive density (ELPD) difference and associated standard errors. We additionally computed stacking weights to quantify the relative predictive contribution of each model.

Posterior Predictive Checks

To validate model fit beyond aggregate predictive accuracy, we conducted posterior predictive checks examining: (1) calibration via Leave-One-Out Probability Integral Transform (LOO-PIT) checks to ensure correspondence between predicted acceptance probabilities and observed acceptance rates); (2) posterior contraction (shrinkage) to verify that data rather than priors drove parameter estimation; (3) sensitivity to stimulus features (correlation between predicted/observed choices and gain/loss); and (4) individual-level fit (correspondence between observed & predicted acceptance rates). These assess whether models capture the key qualitative patterns in the data, independent of formal model comparison metrics.

Results

Simulation

Before presenting empirical results, we validated that the DFT framework can distinguish attentional from valuation mechanisms using choice data alone, through simulation using the empirical stimulus structure ($N = 100$, 200 trials/subject).

Model Recovery We generated synthetic data from either a pure attentional mechanism ($\kappa > 0$, $\lambda = 1$) or a pure valuation mechanism ($\kappa = 0$, $\lambda > 1$). PSIS-LOO correctly identified the generating mechanism in both cases, though with asymmetric sensitivity: when attention was ground truth, M_κ won decisively (ΔELPD = 243.8, SE = 21.6); when valuation was ground truth, M_λ won with a smaller margin (ΔELPD = 9.8, SE = 5.6). This bidirectional recovery rules out a generic model-selection bias toward M_κ.

Parameter Recovery When the correct model was fitted, κ was robustly recovered (Figure 1A); λ recovery was weaker ($r = 0.65$; Panel B), consistent with known identifiability challenges for utility parameters. Crucially, Panel C reveals a systematic confound: when data were generated by attentional bias alone, fitting the misspecified Valuation model yielded

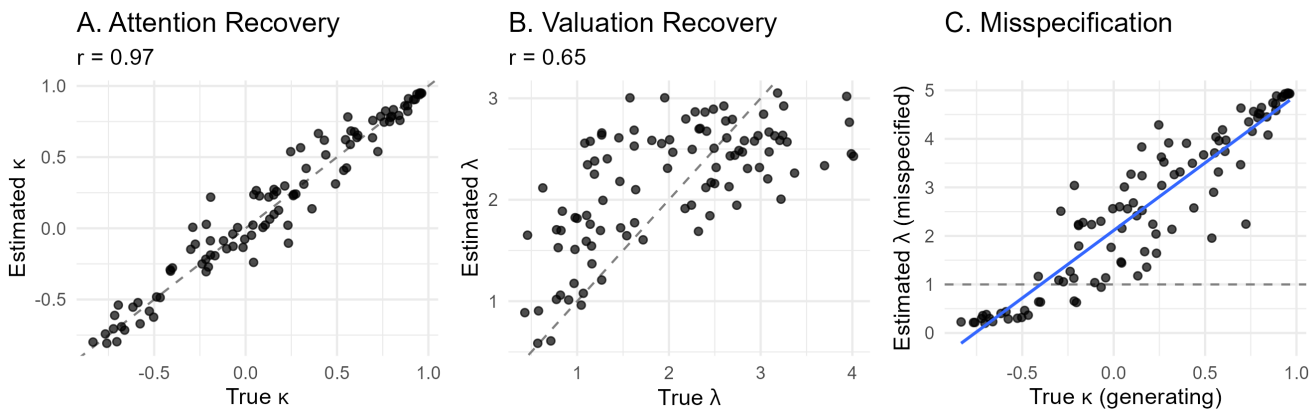

Figure 1: (A) Recovery of κ when data were generated by M_κ ($r = 0.97$). (B) Recovery of λ when data were generated by M_λ ($r = 0.65$). (C) Spurious λ estimates when fitting M_λ to attention-generated data; the linear relationship demonstrates that valuation models can absorb attentional bias into the loss aversion parameter.

spurious loss aversion estimates ($\hat{\lambda} > 1$) that scaled linearly with true κ. This demonstrates that standard utility models may absorb attentional processing into the λ parameter, inflating apparent valuation asymmetry.

Full Model Non-Identifiability We also explored M_{full}, estimating both κ and λ simultaneously. While this yielded marginal improvements in ELPD, posterior estimates exhibited near-perfect negative correlation ($r \approx -0.95$), confirming that choice data alone cannot uniquely locate participants on the continuum between "high valuation / low attention" and "low valuation / high attention." Process-tracing data (e.g., eye-tracking) would be required to resolve this ambiguity.

Model Comparison

Having established identifiability, we compared out-of-sample predictive accuracy across 10 datasets ($N = 686$) using PSIS-LOO.

Figure 2 summarizes the pairwise ΔELPD between M_κ and M_λ. The Attentional model (M_κ) outperformed the Valuation model (M_λ) in 7 out of 10 datasets. In the remaining cases, differences were negligible (confidence intervals overlapping zero), suggesting no decisive advantage for the valuation mechanism. Using the `zhao2020_s1` dataset as a representative example, M_κ provided substantially better fit than M_λ (ΔELPD = 59.6, SE = 10.0), and both models vastly outperformed M_0.

Parameter Estimates

Posterior estimates from the best-fitting M_κ model reveal a consistent attentional bias toward loss outcomes. For instance, in the `zhao2020_s1` dataset, the group-level mean for the attentional weight parameter was $\mu_\kappa = 0.23$ (95% CI $[0.16, 0.30]$). In our parameterization, $\kappa = 0$ implies equal attention ($W_L = 0.5$), while $\kappa = 0.23$ implies that participants allocated approximately 61.5% of their attentional weight to the loss outcome ($W_L = 0.5 + 0.5\kappa$). Table 2 summarizes the posterior estimates of κ across all 10 datasets.

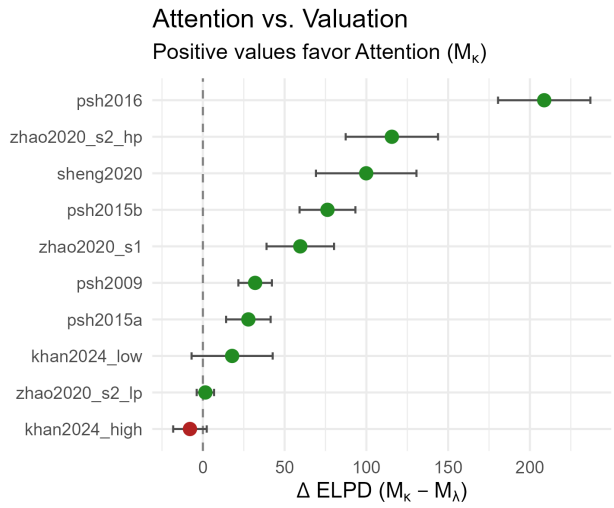

Figure 2: Pairwise difference in expected log pointwise predictive density (ΔELPD) between M_κ and M_λ across 10 datasets. Positive values favor the Attentional model. Error bars represent ±2 SE. Green points indicate datasets where M_κ outperforms M_λ; red points indicate the reverse.

Posterior Predictive Performance

To validate model fit beyond aggregate predictive accuracy, we examined the correspondence between model predictions and observed choice patterns.

Figure 3 displays posterior predictive performance for Zhao et al. (2020) Study 1, selected as a representative case (5th of 10 in relative fit) to avoid cherry-picking. The model achieved 89.6% accuracy (MAE = 0.175; RMSE = 0.285), recovering the indifference boundary through asymmetric attention

Table 2: Model comparison and posterior estimates of κ across datasets (~150–240 trials per participant). Positive ΔELPD favors M_κ over M_λ.

Dataset	N	ΔELPD (SE)	μ_κ [90% CI]
psh2016	95	209.0 (14.1)	0.20 [0.16, 0.23]
zhao2020_s2_hp	52	116.0 (14.1)	0.34 [0.29, 0.38]
sheng2020	93	99.9 (15.4)	0.22 [0.18, 0.26]
psh2015b	47	76.2 (8.5)	0.13 [0.08, 0.18]
zhao2020_s1	43	59.6 (10.3)	0.23 [0.16, 0.30]
psh2009	29	32.0 (5.1)	0.03 [-0.05, 0.10]
psh2015a	34	27.8 (6.8)	0.15 [0.07, 0.23]
khan2024_low	100	17.9 (12.4)	0.02 [-0.01, 0.05]
zhao2020_s2_lp	36	1.5 (2.7)	0.02 [-0.01, 0.05]
khan2024_high	108	–7.9 (5.1)	0.05 [0.04, 0.06]

alone.

Panel A shows that participants exhibit a sharp decision boundary along the diagonal, transitioning rapidly from acceptance (high gain, low loss) to rejection (low gain, high loss). The model successfully captures this topography, though it predicts a slightly more gradual probability gradient across the boundary than observed empirically.

Panel B displays the residual map (Predicted − Observed), revealing that systematic deviations are modest in magnitude (range: $[-0.24, +0.17]$) and confined to specific regions. Two patterns emerge:

The largest negative residual occurs at the symmetric trade-off point (Gain = 10, Loss = 10; residual = −0.24), where observed acceptance rates (48.1%) exceed predictions (23.8%). At this "fair gamble" cell—where gain and loss magnitudes are equal—participants approach indifference, yet the model, with its constant attentional bias toward losses, predicts clear rejection.

The largest positive residual occurs in the high-gain/high-loss region (Gain = 80, Loss = 100; residual = +0.17), where the model predicts 20.4% acceptance but observed rates are only 3.7%. This indicates that extreme magnitudes may trigger additional inhibitory processes not fully captured by linear evidence accumulation with constant κ.

Discussion

The present study examined whether Decision Field Theory can elucidate the mechanisms underlying the tendency to reject gambles that offer positive expected value—a behavioral pattern traditionally attributed to loss-averse valuation (λ). By embedding both valuation asymmetry and attentional asymmetry (κ) within a single hierarchical framework, we tested which mechanism provides a more parsimonious account of rejection behavior.

Across 10 datasets (N = 686; ~140,000 trials), the Attentional model (M_κ) provided better out-of-sample predictions than the Valuation model (M_λ) in seven cases, with no dataset strongly favoring M_λ. Group-level κ estimates ranged from 0.02 to 0.34 (median = 0.14), with 90% credible intervals excluding zero in the majority of datasets—indicating consistent attentional bias toward losses sufficient to produce rejection patterns traditionally attributed to asymmetric utility. Parameter recovery simulations confirmed that these mechanisms are distinguishable from choice data alone, though joint estimation yields non-identifiable posteriors.

Theoretical Implications

These findings—where attention captures variance traditionally attributed to valuation in the majority of datasets—have implications for interpreting λ estimates in the broader literature. Specifically, the superior predictive accuracy of M_κ supports a process account in which rejection of mixed gambles can arise from the over-sampling of loss outcomes during deliberation: because losses enter the evidence stream more frequently ($W_L > 0.5$), accumulated evidence drifts toward rejection even when the utility function is symmetric.

If attentional asymmetry can produce rejection patterns indistinguishable from valuation asymmetry, then elevated λ values may partially reflect attentional processes. Our parameter recovery simulations formalize this concern: fitting a misspecified Valuation model to attention-generated data yielded spurious λ estimates that scaled linearly with true κ (Figure 1C). This suggests that the canonical $\lambda \approx 2$ finding may conflate utility-level and process-level asymmetries, providing one potential explanation for the heterogeneity of λ across elicitation methods (cf. Abburu et al. 2025; Walasek et al. 2021).

The residual analysis reveals boundary conditions of the attentional account. At symmetric trade-offs (Gain = Loss), participants approached indifference despite the model predicting rejection, suggesting an attenuation of loss attention. At high-stakes, the model underpredicted rejection, indicating that large magnitudes may recruit avoidance heuristics beyond constant-κ accumulation. These systematic deviations point toward magnitude-dependent (Mukherjee et al., 2025) extensions—perhaps $\kappa(|x|)$ rather than constant κ—as a direction for future modeling.

Importantly, our findings establish the *sufficiency* of attentional asymmetry to account for rejection behavior, not the exclusion of valuation asymmetry. The non-identifiability of $M_{\lambda\kappa}$ indicates that choice data alone cannot empirically disentangle individuals along the valuation–attention continuum. Rather, we demonstrate that when these mechanisms compete on parsimony grounds, attention provides superior predictive accuracy. Whether DDM-based decompositions (Zhao et al., 2020) and DFT-based decompositions converge on the same individuals remains an open question—specifically, whether participants who exhibit high response bias toward rejection in DDM ($z < 0.5$) are the same participants who exhibit high attentional bias toward losses in DFT ($\kappa > 0$), reflecting a common underlying trait or dissociable processes.

Limitations

While our results suggest that attentional weighting (κ) offers a more parsimonious account of asymmetry in risky choice

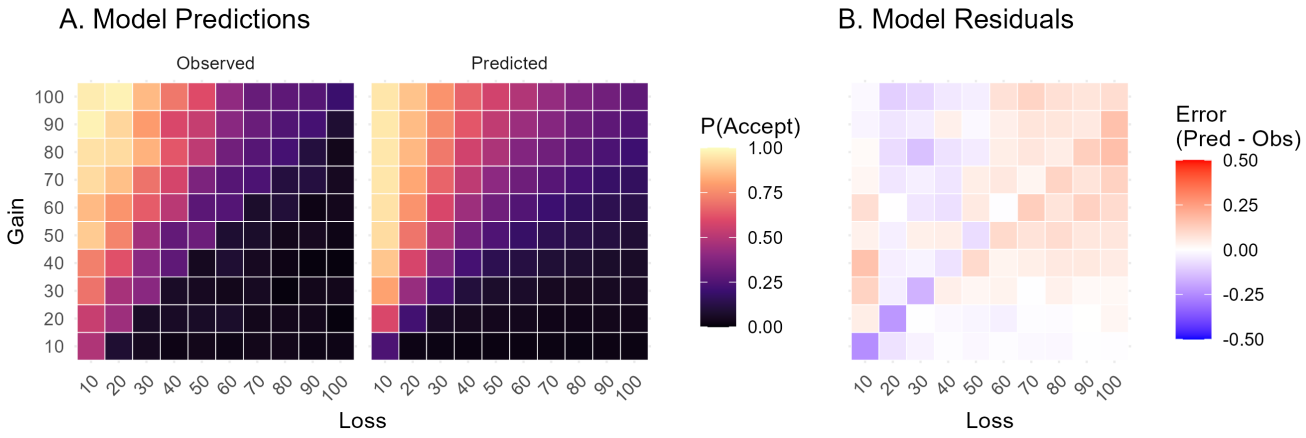

Figure 3: **Posterior predictive performance of the attentional model (M_κ).** (A) Observed and predicted acceptance rates across the gain–loss space in the Zhao et al. (2020) Study 1 dataset. (B) Residual map (Predicted − Observed); blue indicates underestimation, red indicates overestimation. The model achieves 89.6% classification accuracy and recovers the slope of the indifference boundary through asymmetric attention alone ($\kappa = 0.23$; $\lambda = 1$). Systematic residuals are confined to symmetric trade-offs (underestimation at Gain = Loss) and high stakes (overestimation).

than valuation (λ), several constraints should be considered.

First, our reliance on choice outcomes alone leaves temporal dynamics unconstrained. Although parameter recovery simulations confirm that κ is identifiable from choices in this design, the high posterior correlation between λ and κ in M_{full} indicates that these mechanisms remain partially conflated without additional data. Future work should integrate response times or gaze data (Fiedler & Glöckner, 2012; Krajbich et al., 2010) to provide process-level constraints that break this statistical dependency.

Second, we restricted analysis to equiprobable (50/50) gambles to eliminate probability weighting as a confound (effectively making the distraction parameter inert in the DDFT framework; Bhatia (2014)). While this enhances internal validity, it limits generalizability. Whether κ captures similar variance in designs with skewed probabilities (e.g., rare-event paradigms) remains an open question.

Third, our DFT implementation assumes constant drift, capturing the *net* attentional bias rather than dynamic attention switching as modeled in full multi-attribute DFT (Diederich, 1997). Relatedly, our residual analysis revealed systematic deviations at magnitude extremes (Figure 3B), suggesting that a constant-κ assumption may not hold across the full stimulus range. Extensions incorporating magnitude-dependent attention or threshold-based heuristics warrant exploration.

Fourth, our DFT implementation does not include asymmetric decision thresholds, which would capture response bias toward rejection. Prior DDM work has identified starting-point bias as a significant contributor to loss-averse behavior (Sheng et al., 2020; Zhao et al., 2020). Whether DFT's κ and DDM's z capture the same underlying process or represent distinct mechanisms requires future cross-model comparison.

Finally, we consider the challenge of model mimicry (Stott, 2006). While our recovery simulations demonstrate that κ and λ are mathematically distinct within this framework, it remains possible that the attentional parameter κ absorbs variance from unmodeled processes, such as nonlinearities in the comparison stage or response biases. The current findings, however, suggest that the behavioral phenomenon traditionally mapped onto a valuation parameter is mathematically better described as a bias in evidence accumulation.

Conclusion

We developed a hybrid Decision Field Theory framework that embeds PT-style valuation asymmetry (λ) into DFT's structure to adjudicate between valuation and attentional accounts of mixed-gamble rejection using binary choice data.

Across 10 datasets, attentional asymmetry (κ) provided superior predictive accuracy, with participants allocating approximately 57–67% of processing weight to losses. Simulations confirmed these mechanisms are distinguishable via DFT's endogenous variance, though joint estimation requires process-tracing data. Fitting misspecified valuation models to attention-generated data produced spurious λ estimates, suggesting standard utility models absorb attentional variance into inflated loss-aversion parameters.

These findings demonstrate that attentional bias is *sufficient* to account for rejection behavior within this framework, providing an alternative generative explanation to asymmetric valuation. More broadly, this work represents one attempt to augment process models with constructs from descriptive theories (here, λ from Prospect Theory), an approach that may prove useful for adjudicating between mechanisms that neither framework could easily distinguish alone.

Acknowledgments

We thank the reviewers and the meta-reviewer for their helpful comments. AI models (Opus 4.5/GPT 5.2 primarily) were

used in validation and optimisation of the modeling pipeline.

References

Abburu, A., Agarwal, S., & Mukherjee, S. (2025). Heterogeneity in Loss Aversion Estimates across Modeling Approaches. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*(0). https://escholarship.org/uc/item/24w9p95m

Bhatia, S. (2014). Sequential sampling and paradoxes of risky choice. *Psychonomic Bulletin & Review*, *21*(5), 1095–1111. https://doi.org/10.3758/s13423-014-0650-1

Brown, A. L., Imai, T., Vieider, F., & Camerer, C. F. (2023). Meta-analysis of empirical estimates of loss aversion. *Journal of Economic Literature*. https://doi.org/10.1257/jel.20221698

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Clay, S. N., Clithero, J. A., Harris, A. M., & Reed, C. L. (2017). Loss Aversion Reflects Information Accumulation, Not Bias: A Drift-Diffusion Model Study. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01708

Diederich, A. (1997). Dynamic Stochastic Models for Decision Making under Time Constraints. *Journal of Mathematical Psychology*, *41*(3), 260–274. https://doi.org/10.1006/jmps.1997.1167

Fiedler, S., & Glöckner, A. (2012). The Dynamics of Decision Making in Risky Choice: An Eye-Tracking Analysis. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00335

Gabry, J., Češnovar, R., Johnson, A., & Bronder, S. (2025). *Cmdstanr: R Interface to 'CmdStan'*. manual. https://mc-stan.org/cmdstanr/

Gal, D., & Rucker, D. D. (2018). The Loss of Loss Aversion: Will It Loom Larger Than Its Gain? (S. Shavitt, Ed.). *Journal of Consumer Psychology*, *28*(3), 497–516. https://doi.org/10.1002/jcpy.1047

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kellen, D., Steiner, M. D., Davis-Stober, C. P., & Pappas, N. R. (2020). Modeling choice paradoxes under risk: From prospect theories to sampling-based accounts. *Cognitive Psychology*, *118*, 101258. https://doi.org/10.1016/j.cogpsych.2019.101258

Khan, O., Mukherjee, S., Srinivasan, N., & Abburu, A. (2024). *Role of Magnitude in Loss Aversion*. OSF. https://doi.org/10.17605/OSF.IO/4CHP2

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635

Lejarraga, T., Schulte-Mecklenbeck, M., Pachur, T., & Hertwig, R. (2019). The attention–aversion gap: How allocation of attention relates to loss aversion. *Evolution and Human Behavior*, *40*(5), 457–469. https://doi.org/10.1016/j.evolhumbehav.2019.05.008

Mrkva, K., Johnson, E. J., Gächter, S., & Herrmann, A. (2020). Moderating Loss Aversion: Loss Aversion Has Moderators, But Reports of its Death are Greatly Exaggerated. *Journal of Consumer Psychology*, *30*(3), 407–428. https://doi.org/10.1002/jcpy.1156

Mukherjee, S., Khan, O., & Srinivasan, N. (2025). The role of magnitude in loss aversion. *Decision*, *12*(2), 93–110. https://doi.org/10.1037/dec0000256

Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology. General*, *147*(2), 147–169. https://doi.org/10.1037/xge0000406

Pachur, T., & Zilker, V. (2023). Psychological Theories of Decision Making Under Risk. In G. Pogrebna & T. Hills (Eds.), *The Cambridge handbook of behavioral data science*. Cambridge University Press.

Papaspiliopoulos, O., Roberts, G. O., & Sköld, M. (2007). A General Framework for the Parametrization of Hierarchical Models. *Statistical Science*, *22*(1). https://doi.org/10.1214/088342307000000014

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1446–1465. https://doi.org/10.1037/a0013646

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., Davison, C., Demić, E., Esteban-Serna, C., Friedemann, M., Gibson, S. P., Jarke, H., Karakasheva, R., Khorrami, P. R., Kveder, J., Andersen, T. L., Lofthus, I. S., McGill, L., Nieto, A. E., … Folke, T. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, *4*(6), 622–633. https://doi.org/10.1038/s41562-020-0886-x

Sheng, F., Ramakrishnan, A., Seok, D., Zhao, W. J., Thelaus, S., Cen, P., & Platt, M. L. (2020). Decomposing loss aversion from gaze allocation and pupil dilation. *Proceedings of the National Academy of Sciences*, *117*(21), 11356–11363. https://doi.org/10.1073/pnas.1919670117

Sokol-Hessner, P., Hartley, C. A., Hamilton, J. R., & Phelps, E. A. (2015). Interoceptive ability predicts aversion to losses. *Cognition & Emotion*, *29*(4), 695–701. https://doi.org/10.1080/02699931.2014.925426

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. https://doi.org/10.1073/pnas.0806761106

Sokol-Hessner, P., Lackovic, S. F., Tobe, R. H., Camerer, C. F., Leventhal, B. L., & Phelps, E. A. (2015). Determinants of Propranolol's Selective Effect on Loss Aversion. *Psychological Science*, *26*(7), 1123–1130. https://doi.org/10.1177/0956797615582026

Sokol-Hessner, P., Raio, C. M., Gottesman, S. P., Lackovic, S. F., & Phelps, E. A. (2016). Acute stress does not affect risky monetary decision-making. *Neurobiology of Stress*, *5*, 19–25. https://doi.org/10.1016/j.ynstr.2016.10.003

Spektor, M. S., Kellen, D., Rieskamp, J., & Klauer, K. C. (2024). Absolute and relative stability of loss aversion across contexts. *Journal of Experimental Psychology: General*, *153*(2), 454–472. https://doi.org/10.1037/xge0001513

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32*(2), 101–130. https://doi.org/10.1007/s11166-006-8289-6

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Walasek, L., Mullett, T. L., & Stewart, N. (2021). Acceptance of mixed gambles is sensitive to the range of gains and losses experienced, and estimates of lambda (λ) are not a reliable measure of loss aversion: Reply to André and de Langhe (2021). *Journal of Experimental Psychology: General*, *150*(12), 2666–2670. https://doi.org/10.1037/xge0001054

Walasek, L., & Stewart, N. (2015). How to make loss aversion disappear and reverse: Tests of the decision by sampling origin of loss aversion. *Journal of Experimental Psychology: General*, *144*(1), 7–11. https://doi.org/10.1037/xge0000039

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, *139*(2), 497–518. https://doi.org/10.1037/a0029383

Zhao, W. J., Walasek, L., & Bhatia, S. (2020). Psychological mechanisms of loss aversion: A drift-diffusion decomposition. *Cognitive Psychology*, *123*, 101331. https://doi.org/10.1016/j.cogpsych.2020.101331