# Probability Weighting from Intrapersonal Aggregation: Internal Sampling, Metacognition, and Divergence Minimization

**Stephan Hartmann (S.Hartmann@lmu.de)**[1,2]
[1]Munich Center for Mathematical Philosophy, LMU Munich
[2]Department of Psychology, LMU Munich

## Abstract

Probability weighting is a central component of cumulative prospect theory, but its cognitive basis remains contested. We propose a process-level account in which weighting emerges from *intrapersonal aggregation*: decision makers integrate an externally provided probability with internally generated estimates from memory or simulation. We formalize this as minimization of a weighted Kullback–Leibler divergence and show that the solution is logarithmic pooling in log-odds space, yielding the standard two-parameter weighting function. Curvature reflects the relative influence of external information, while elevation captures directional bias in internal sampling. The model predicts that sampling depth and external reliability primarily affect curvature, whereas framing or affective biases primarily shift elevation. An analytic Beta-sampling benchmark makes these dependencies explicit.

**Keywords:** prospect theory; probability weighting; internal sampling; metacognition; belief aggregation; KL divergence

## Introduction

Probability weighting—the systematic nonlinear transformation of objective probabilities—is a central component of prospect theory and its cumulative extension (CPT) (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Empirically, subjective probabilities are well captured by an inverse-S-shaped weighting function: small probabilities are overweighted, large probabilities underweighted, and intermediate probabilities transformed smoothly. This pattern is remarkably stable across elicitation methods, task domains, and populations, and is well described by simple two-parameter functional forms whose estimates cluster in relatively narrow ranges (Gonzalez & Wu, 1999; Wakker, 2010). The robustness and regularity of the phenomenon point to an underlying mechanism rather than unsystematic error.

Despite this progress, there is no consensus account of why probability weighting takes this particular functional form and why its parameters vary systematically across contexts. Recent work has linked nonlinear weighting to Bayesian prior integration, noisy or finite-precision representations, linear transformations in log-odds space, and attentional allocation (Fennell & Baddeley, 2012; Juechems et al., 2021; Pachur et al., 2018; Zhang & Maloney, 2012; Zhang et al., 2020; Zilker & Pachur, 2022).

These accounts provide process-level explanations of probability distortion, but they do not yet fully explain the stability of the CPT functional form together with the distinct roles of its parameters. The present paper addresses this gap by asking whether the standard two-parameter weighting function can be derived from a general principle for integrating externally provided probabilities with internally generated estimates.

A prominent class of psychological explanations appeals to *internal sampling*. According to sampling-based accounts, people's judgments may reflect samples from internal probability distributions rather than deterministic best estimates (Costello & Watts, 2014; Vul & Pashler, 2014). Vul and Pashler, for example, show that averaging two estimates from the same individual can improve accuracy, suggesting partially independent internal samples. Such models explain variability and regression effects in probabilistic judgment. However, sampling alone does not explain why subjective probabilities should be transformed by a stable inverse-S-shaped function, nor why this transformation should take a consistent parametric form. Additional structure is required to explain how internally generated samples are combined into a single, coherent probability judgment.

The closest precursor is Fennell and Baddeley (2012), who argue that nonlinear weighting can arise when stated probabilities are combined with prior assumptions derived from experience. We share this anti-primitivist perspective but differ in the formal mechanism. Whereas Fennell and Baddeley treat the prior as a distribution over objective probabilities, we treat internal sampling as generating subjective probability estimates; the relevant normative structure is therefore intrapersonal aggregation rather than Bayesian updating.

Our central idea is that probability weighting arises from *intrapersonal aggregation*: the confidence-sensitive integration of an externally provided probability with internally generated probability estimates. We formalize this integration problem as a principled aggregation task. The agent selects a single subjective probability that best represents the available information by minimizing a weighted average Kullback–Leibler (KL) divergence between the aggregate probability and each source. For binary events, this divergence-minimization problem is equivalent to logarithmic opinion pooling. Our key result is that the solution of this intrapersonal aggregation problem yields the standard two-parameter form of probability weighting used in CPT.

This derivation assigns clear roles to the CPT parameters: curvature reflects the balance between internal sampling and external information, while elevation captures systematic bias in internal sampling. This distinction parallels Wakker's sep-



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

aration of likelihood sensitivity from pessimism/optimism as two components of probabilistic risk attitude (Wakker, 2010). On this view, probability weighting is not a primitive distortion of probability, but the outcome of a cognitively plausible integration problem under bounded internal sampling and limited metacognitive access.

Beyond the formal derivation, the model yields sharp and testable predictions. Manipulations that increase effective internal sampling or confidence in external information should selectively attenuate curvature, whereas manipulations that bias internal sampling—such as framing or affective influences—should selectively shift elevation. These dissociations go beyond curve fitting and specify how distinct cognitive processes map onto distinct parameters of the weighting function.

The contribution of this paper is therefore not to fit data, but to show that the canonical CPT weighting function can be derived from a simple and cognitively plausible integration principle, which in turn yields parameter-specific empirical predictions.

## Cognitive motivation: internal sampling and metacognitive integration

The cognitive plausibility of the model rests on two assumptions: people generate sparse internal samples when judging probabilities, and they weight available representations by imperfect confidence signals. Together, they suggest that probability judgment integrates multiple imperfect representations generated within a single cognitive system.

A large body of evidence indicates that people often rely on *internal sampling* when making judgments under uncertainty. Rather than always reporting deterministic best estimates, individuals may generate responses that behave like samples from internal probability distributions (Costello & Watts, 2014; Vul & Pashler, 2014). These samples may correspond to recalled instances, imagined scenarios, or stochastic mental simulations of outcomes. Sampling-based models explain variability, regression toward the mean, and sensitivity to time pressure or cognitive load. Related Bayesian sampling models show how generic sampling-based inference can produce systematic distortions and incoherence in probability judgment (Zhu et al., 2020). Crucially, internal sampling is typically sparse: people generate only a few samples, and these samples are neither perfectly independent nor unbiased, reflecting constraints on attention, working memory, and deliberation time.

Sampling-based accounts, however, leave open why subjective probabilities should be transformed by a stable inverse-S-shaped function across tasks and elicitation methods. If judgments were simple averages of a small number of noisy samples, distortions would vary idiosyncratically with sample size or noise level. The empirical regularity of probability weighting therefore suggests that internally generated samples are combined according to a stable aggregation principle.

This observation aligns with work on the "wisdom of the inner crowd" (Herzog & Hertwig, 2014). When individuals generate multiple estimates of the same quantity—especially under conditions that reduce redundancy—aggregating these estimates can improve accuracy through error cancellation. These findings support the idea that people have access to multiple internal probabilistic representations and that these can be fruitfully combined, but they do not specify how such combination proceeds when explicit external probabilities are also available.

A second ingredient is *metacognition*: the capacity to monitor, albeit imperfectly, the reliability of one's own representations. Confidence signals are systematically related to evidence accumulation, decision time, and time pressure (Pleskac & Busemeyer, 2010), and provide a plausible basis for implicitly weighting internally generated and externally provided representations.

Metacognitive integration becomes particularly salient when explicit external probabilistic information is available. In many decision tasks, individuals are presented with base rates, frequencies, or expert forecasts while simultaneously generating internal estimates based on experience or imagination. These sources are not treated as interchangeable: external probabilities are weighted more heavily when perceived as reliable, and internal judgments more strongly when external sources are uncertain. This pattern suggests confidence-sensitive integration rather than simple substitution or averaging.

Taken together, these considerations motivate a specific computational problem. At the time of judgment, a decision maker has access to an externally provided probability and to a small set of internally generated probability estimates that are noisy, partially redundant, and potentially biased. The task is to integrate these representations into a single, coherent subjective probability. The central claim of this paper is that this integration problem is naturally modeled as a form of *intrapersonal belief aggregation*. In the next section, we formalize this idea by modeling integration as the minimization of a weighted divergence between a candidate aggregate probability and the set of internal and external estimates.

## Model: divergence-minimizing intrapersonal aggregation

We now formalize the cognitive picture outlined above. The guiding idea is to treat probability judgment as a problem of *intrapersonal aggregation*: at the moment of judgment, a single agent has access to multiple probabilistic representations of the same event and must integrate them into one coherent subjective probability suitable for guiding choice.

Consider a binary event $E$. Let $p \in (0, 1)$ denote an externally provided probability for $E$, such as a base rate, frequency, or expert forecast. In addition, the agent generates a small set of internally produced probability estimates $p_1, \ldots, p_n \in (0, 1)$ through memory retrieval, imagination, or mental simulation. These internal estimates reflect bounded internal sampling: they are noisy, partially redundant, and potentially biased. The task is to combine these representations into a single subjective probability $p'$.

We represent each probability as a Bernoulli distribution. Let $P = (p, 1-p)$ denote the external distribution, let $P_i = (p_i, 1-p_i)$ for $i = 1, \ldots, n$ denote the internal distributions, and let $P' = (p', 1-p')$ denote the aggregate distribution to be determined. The central modeling assumption is that the agent selects $P'$ so as to best represent the available distributions as a whole, taking into account their relative epistemic status.

We model this integration problem as the minimization of a weighted average Kullback–Leibler (KL) divergence:

$$L(p') = \gamma\, D_{\mathrm{KL}}(P' \| P) + \frac{1-\gamma}{n} \sum_{i=1}^{n} D_{\mathrm{KL}}(P' \| P_i). \quad (1)$$

The parameter $\gamma \in [0, 1]$ captures the relative influence of the external probability compared to the internally generated estimates. Cognitively, $\gamma$ reflects confidence in the external source, or more generally the degree to which externally provided information is privileged over internal sampling.

The interpretation of Eq. (1) is straightforward. KL divergence measures the expected log-loss incurred when one distribution is used in place of another. Minimizing a weighted average of KL divergences therefore selects an aggregate probability that is, in an information-theoretic sense, as close as possible to each available estimate, while allowing more trusted sources to exert greater influence. This formulation specifies a computational-level solution to the integration problem; it does not assume that agents explicitly compute divergences or solve optimization problems.

The choice of KL divergence is substantive. It treats $P'$ as the distribution from which the agent will act and evaluates the expected log-loss of using it in place of each available representation. Minimizing Eq. (1) therefore balances fidelity to multiple sources under a proper scoring rule. This differs from linear pooling, which averages probabilities directly, and from max or heuristic pooling, which discards information from less salient sources. Logarithmic pooling is appropriate here because the internal estimates are not independent observations for a single Bayesian update, but partially redundant probabilistic representations whose credibility must be balanced. It also connects to log-odds accounts of probability representation (Zhang & Maloney, 2012; Zhang et al., 2020).

Alternative pooling rules capture different cognitive strategies but do not serve the same role. Linear pooling averages probabilities directly, but it does not preserve the log-odds structure identified in recent representational accounts and lacks the dynamic coherence properties associated with logarithmic pooling. Max pooling or winner-take-all rules may be plausible under severe attentional constraints, but they discard information from non-selected sources. Bayesian updating would be appropriate if the internal estimates were independent pieces of evidence generated by a known likelihood model; here they are better understood as partially redundant internal representations whose credibility must be balanced.

In the belief aggregation literature, minimizing average KL divergence is equivalent to *logarithmic opinion pooling*, an aggregation rule with well-known coherence properties (Dietrich & List, 2014; Pettigrew & Weisberg, forthcoming). In the present context, the pooling is intrapersonal rather than interpersonal: the "opinions" being aggregated are distinct probabilistic representations generated within a single cognitive system. (For related uses of divergence-based and Bayesian frameworks in modeling reasoning, see Eva and Hartmann (2018) and Eva et al. (2020).)

Minimizing the objective in Eq. (1) yields a particularly simple solution when expressed in log-odds space. The aggregate log-odds are a convex combination of the external log-odds and the average internal log-odds:

$$\log \frac{p'}{1-p'} = \gamma \log \frac{p}{1-p} + \frac{1-\gamma}{n} \sum_{i=1}^{n} \log \frac{p_i}{1-p_i}. \quad (2)$$

Thus, the aggregation rule amounts to confidence-weighted averaging of log-odds rather than of probabilities. This feature is crucial: it ensures coherence and leads directly to the functional form of probability weighting. It also provides a derivation of the linear log-odds structure that Zhang and Maloney (2012) identify descriptively across perception, memory, and decision-making.

Solving Eq. (2) for $p'$ yields the following result.

**Proposition 1** (Probability weighting from intrapersonal aggregation)**.** *Let $P$ and $P_1, \ldots, P_n$ be Bernoulli distributions with parameters $p$ and $p_1, \ldots, p_n$, and let $P'$ minimize Eq.* (1)*. Then the resulting aggregate probability $p'$ is given by*

$$p' = w(p) = \frac{\delta\, p^{\gamma}}{\delta\, p^{\gamma} + (1-p)^{\gamma}}, \quad (3)$$

*where*

$$\delta := \left( \prod_{i=1}^{n} \frac{p_i}{1-p_i} \right)^{\frac{1-\gamma}{n}}. \quad (4)$$

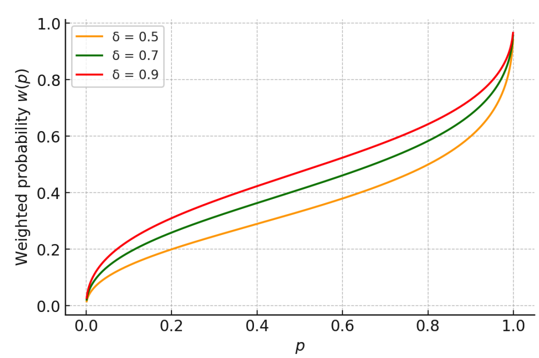


Figure 1: Probability weighting function from eq. (3) for $\gamma = 0.5$ and varying $\delta$ values ($\delta = 0.5,\ 0.7,\ 0.9$). The function exhibits the typical inverse-S curvature observed in empirical studies.

Two features of the result are central. First, the exponent $\gamma$ directly reflects the relative weight assigned to external versus internal information in the aggregation objective. Greater reliance on internal sampling (lower $\gamma$) produces stronger nonlinearity, while greater confidence in the external source (higher $\gamma$) yields more linear weighting. Second, the elevation parameter $\delta$ summarizes the geometric mean of internal odds and therefore captures directional bias in internal sampling. If internal estimates are, on average, unbiased in log-odds space, then $\delta \approx 1$, yielding a weighting function symmetric around the diagonal. Systematic optimism or pessimism in internal sampling shifts $\delta$ above or below unity without altering the basic shape (see also Figure 1).

The model thus derives the CPT weighting function from a general aggregation problem rather than postulating it as a primitive transformation. It shows how sparse, noisy internal estimates and externally provided probabilities can be integrated under confidence-sensitive weighting and coherence constraints. In the next section, we use this structure to derive empirical predictions linking manipulations of internal sampling, bias, and confidence to selective changes in curvature and elevation.

## Predictions

A key advantage of the divergence-minimizing aggregation model is that it links cognitive manipulations to specific parameters of the probability weighting function. Rather than treating curvature and elevation as free descriptive parameters, the model assigns them distinct roles: the exponent $\gamma$ captures the relative influence of internal sampling versus external information, whereas the elevation parameter $\delta$ captures directional bias in internal sampling via the geometric mean of internal odds. This yields structured, testable predictions about how different manipulations should map onto parameter changes.

These predictions are continuous with existing findings linking probability weighting to representational precision, attention allocation, and information search, but they impose a stronger constraint: manipulations of precision or source reliability should primarily affect curvature, whereas manipulations that bias the content of internal sampling should primarily affect elevation. Importantly, the model predicts *dominant* rather than strictly exclusive effects. Since $\delta$ depends both on the distribution of internal odds and on the weight $(1 - \gamma)$, secondary interactions between curvature and elevation are expected.

**Sampling depth and independence.** Manipulations that increase the effective depth or independence of internal sampling, such as extended deliberation, repeated elicitation, or structured reconsideration, should primarily affect curvature. Greater sampling depth stabilizes internal log-odds representations and should therefore attenuate inverse-S-shaped distortion, yielding fitted curvature parameters $\hat{\gamma}$ closer to unity. Conversely, time pressure or cognitive load should reduce effective sampling depth and amplify curvature effects. This prediction is consistent with accounts that link probability distortion to representational precision, resource constraints, and systematic changes in log-odds slope with sample numerosity (Juechems et al., 2021; Zhang & Maloney, 2012).

Provided that such manipulations do not systematically alter the *mean* of internal log-odds, their effect on elevation $\delta$ should be limited. Indirect effects are nevertheless possible if changes in sampling depth also affect the distribution of internal estimates, for example by changing redundancy, variability, or calibration. The model thus predicts a dominant effect on curvature, with at most secondary effects on elevation.

**Directional bias in internal sampling.** Manipulations that induce systematic bias in internally generated estimates should primarily affect elevation. Affective states, motivational factors, or framing manipulations can skew memory retrieval or mental simulation in an optimistic or pessimistic direction. In the model, such bias operates in log-odds space and enters through the geometric mean of internal odds, producing a vertical shift of the weighting function.

The prediction is a systematic shift in $\hat{\delta}$ with comparatively smaller changes in curvature $\hat{\gamma}$, yielding a dissociation between bias and probability sensitivity. This is in line with attentional and sampling-based accounts on which systematic biases in information acquisition influence probability weighting (Pachur et al., 2018; Zilker & Pachur, 2022). The present model adds the more specific claim that such effects should primarily register in elevation rather than in overall curvature.

**Confidence in external information.** Manipulations that alter the perceived reliability of the external probability should primarily modulate curvature by changing the relative weight assigned to the external source. Increasing perceived credibility, for example via expertise cues, precision indicators, or accuracy incentives, should increase $\gamma$ and reduce nonlinearity, whereas decreasing trust should have the opposite effect. This connects to prior-integration accounts in which stated probabilities are combined with background expectations whose influence depends on perceived reliability (Fennell & Baddeley, 2012).

Because $\delta$ is scaled by $(1 - \gamma)$, increasing $\gamma$ also attenuates the influence of internal odds on the aggregate probability. As external information becomes dominant, elevation effects should therefore be pulled toward neutrality, that is, toward $\delta = 1$. The model thus predicts a primary effect on curvature together with a secondary attenuation of elevation.

**A minimal experimental test.** These predictions can be tested within a compact within-subject design. Participants make probability judgments or risky choices based on explicit probabilities while external-source reliability and internal sampling depth are independently manipulated. An additional manipulation can bias internal sampling, for example via framing or affect.

The model predicts parameter-specific shifts: reliability and sampling-depth manipulations should primarily shift fitted curvature $\hat{\gamma}$, whereas bias manipulations should primarily shift elevation $\hat{\delta}$. Empirically, these parameters can be estimated by fitting standard CPT weighting functions, such as the Gonzalez–Wu form, across a probability grid. Parameter recovery is approximate and depends on the fitting procedure, but the critical prediction concerns systematic differences across conditions rather than absolute parameter values.

A distinctive implication of the model is that curvature and elevation are *partially dissociable* under targeted manipulations, in contrast to accounts in which a single noise, precision, or attention parameter drives probability weighting as a whole. Because the aggregation rule is fixed, the functional form of probability weighting should remain stable across conditions, with variability expressed primarily through structured parameter shifts rather than changes in overall shape.

## Analytic benchmark and illustrative checks

This section introduces a compact analytic benchmark that makes the dependence of the elevation parameter on properties of internal sampling explicit. The aim is diagnostic clarity rather than simulation-based model evaluation. The primary contribution of the paper is the theoretical derivation of the weighting function and the analytic interpretation of its parameters.

**Analytic benchmark: Beta-distributed internal samples.** To obtain a tractable characterization of elevation, suppose that internal probability estimates $p_1, \dots, p_n$ are independent draws from a Beta distribution, $p_i \sim \mathrm{Beta}(\alpha, \beta)$. This assumption is not intended as a literal cognitive model, but as a convenient parametric family that allows central tendency and asymmetry of internal sampling to be varied while holding the aggregation rule fixed. In this respect, it is related to Bayesian prior-based accounts such as Fennell and Baddeley (2012), but serves a different role: here the Beta distribution characterizes the distribution of internally generated estimates rather than an explicit prior over objective probabilities.

Recall that elevation is given by the geometric mean of internal odds, $\delta = \left(\prod_{i=1}^{n} \frac{p_i}{1-p_i}\right)^{\frac{1-\gamma}{n}}$. Under Beta sampling, its expectation can be computed in closed form.

**Proposition 2** (Exact expectation under Beta sampling)**.** *Let* $p_i \overset{i.i.d.}{\sim} \mathrm{Beta}(\alpha, \beta)$. *For any finite n,*

$$\mathbb{E}[\delta_n] = \left[\frac{\mathrm{B}\left(\alpha + \frac{1-\gamma}{n}, \beta - \frac{1-\gamma}{n}\right)}{\mathrm{B}(\alpha, \beta)}\right]^n .$$

As the number of internal samples grows, $\delta_n$ converges almost surely to a simple expression involving the digamma function.

**Proposition 3** (Almost sure limit)**.** *As* $n \to \infty$,

$$\delta_n \xrightarrow{a.s.} \exp\left\{(1-\gamma)\left[\psi(\alpha) - \psi(\beta)\right]\right\} .$$

| Regime | Params | $\delta$ | Interpretation |
|---|---|---|---|
| Bal. | $\alpha = \beta$ | $\approx 1$ | no bias |
| Low | $\alpha < \beta$ | $< 1$ | pessimistic |
| High | $\alpha > \beta$ | $> 1$ | optimistic |
| Inf. | $\alpha = pk,\ \beta = (1-p)k$ | varies | tracks $p$ |

Table 1: Qualitative effects of internal sampling.

These results make the interpretation of elevation transparent. Asymmetry in the internal sampling distribution ($\alpha \neq \beta$) induces a stable shift in elevation, while symmetric sampling yields $\delta = 1$. In particular, low-biased internal sampling ($\alpha < \beta$) yields $\delta < 1$, whereas high-biased internal sampling ($\alpha > \beta$) yields $\delta > 1$. Curvature, by contrast, is governed by $\gamma$.

**Illustrative checks.** The Beta benchmark also clarifies the qualitative patterns expected under representative internal sampling regimes. Table 1 summarizes the resulting parameter shifts.

A useful contrast is provided by an *informed* internal sampling regime, in which the Beta distribution is centered on the externally given probability itself, i.e. $\alpha = pk$ and $\beta = (1-p)k$ for some $k > 0$. In this case, internally generated estimates track the external probability across the probability grid, and the resulting weighting pattern can differ qualitatively from the usual inverse-S shape. This contrast is diagnostic: inverse-S-shaped weighting is not a generic consequence of noise or limited sampling, but depends on the structure and bias of internally generated estimates.

Together, the analytic benchmark and these illustrative checks show how divergence-minimizing intrapersonal aggregation yields structured and interpretable weighting patterns. Their role is not to provide a full empirical validation of the model, but to clarify how the theoretical mechanism maps onto observable parameter shifts and to identify concrete targets for future empirical tests.

## Discussion

This paper has proposed a process-level account of probability weighting that connects internal sampling, metacognitive confidence, and principled belief aggregation. The model derives the standard two-parameter weighting function from a cognitively plausible integration problem: combining sparse, noisy internal estimates with externally provided probabilities under confidence-sensitive weighting. The result is a direct mapping from a well-defined aggregation rule to the functional form used in cumulative prospect theory.

The account complements existing process explanations. Bayesian models explain weighting as the integration of stated probabilities with priors (Fennell & Baddeley, 2012); finite-precision and log-odds accounts emphasize noisy, resource-limited, or transformed representations (Juechems et al., 2021; Zhang & Maloney, 2012; Zhang et al., 2020); and attentional accounts link CPT parameters to information search (Pachur et al., 2018; Zilker & Pachur, 2022). The present contribution is to locate probability weighting at the level of integration: externally provided probabilities are combined with internally generated estimates by confidence-sensitive aggregation.

A central contribution of the account is the interpretability of the weighting parameters. The model assigns distinct roles to curvature and elevation and explains how they arise from separate features of the underlying cognitive process. The analytic benchmark sharpens this distinction by showing how asymmetries in internal sampling induce stable shifts in elevation without affecting curvature. This separation yields clear empirical dissociations that go beyond descriptive curve fitting and provide a mechanism-level interpretation of parameter variation.

The model also reframes the normative status of probability weighting. From the perspective of expected utility theory, nonlinear weighting is often interpreted as an error. On the present account, however, inverse-S-shaped weighting can arise as a rational outcome of bounded information integration. When internal evidence is sparse, partially redundant, and imperfectly calibrated relative to external information, divergence-minimizing aggregation yields systematic nonlinearity without presupposing mistake. Probability weighting thus reflects how bounded agents integrate multiple imperfect sources of information under uncertainty.

The account has clear limitations. The formal model is restricted to binary events, leaving extensions to multinomial outcomes and full cumulative weighting for future work. Internal sampling is treated abstractly, without specifying the cognitive or neural mechanisms by which samples are generated or correlated over time. Moreover, the model is not yet directly fitted to behavioral data and does not provide an algorithmic-level implementation of the aggregation process. Its primary role is to provide a computational-level derivation and parameter interpretation that can be tested and embedded within richer process models.

More broadly, the framework illustrates the potential of intrapersonal belief aggregation as a modeling strategy in cognitive science. Many cognitive tasks involve integrating multiple internal representations rather than aggregating information across individuals. Applying aggregation theory at the intrapersonal level allows normative principles to inform cognitive models without collapsing into idealized rationality assumptions. In the present case, this perspective yields a compact explanation of probability weighting and generates precise, testable predictions for future experimental work.

Future work should test the predicted parameter dissociations in targeted experiments, relate the model more directly to behavioral data, and extend the framework to multinomial events, dynamic learning contexts, and algorithmic implementations.

## Appendix: Proofs and technical details

This appendix provides concise proof sketches for the main formal results. Details are kept to a minimum to conserve space; all steps rely on standard properties of KL divergence and the Beta distribution.

### Proof of Proposition 2

Let $p_i \overset{\text{i.i.d.}}{\sim} \mathrm{Beta}(\alpha, \beta)$ and define

$$\delta_n = \prod_{i=1}^{n} \left( \frac{p_i}{1-p_i} \right)^{\frac{1-\gamma}{n}}.$$

By independence,

$$\mathbb{E}[\delta_n] = \left( \mathbb{E}\left[ \left( \frac{p}{1-p} \right)^{\frac{1-\gamma}{n}} \right] \right)^n, \quad p \sim \mathrm{Beta}(\alpha, \beta).$$

Using the Beta density,

$$\mathbb{E}\left[ \left( \frac{p}{1-p} \right)^t \right] = \frac{1}{\mathrm{B}(\alpha,\beta)} \int_0^1 p^{\alpha+t-1} (1-p)^{\beta-t-1}\, dp \tag{5}$$

$$= \frac{\mathrm{B}(\alpha+t, \beta-t)}{\mathrm{B}(\alpha,\beta)}, \tag{6}$$

for $\alpha + t > 0$ and $\beta - t > 0$. Substituting $t = (1-\gamma)/n$ yields the result. □

### Proof of Proposition 3

Taking logs,

$$\log \delta_n = \frac{1-\gamma}{n} \sum_{i=1}^{n} \log \frac{p_i}{1-p_i}.$$

By the strong law of large numbers,

$$\frac{1}{n} \sum_{i=1}^{n} \log \frac{p_i}{1-p_i} \xrightarrow{\text{a.s.}} \mathbb{E}\left[ \log \frac{p}{1-p} \right], \quad p \sim \mathrm{Beta}(\alpha, \beta).$$

For the Beta distribution,

$$\mathbb{E}[\log p] = \psi(\alpha) - \psi(\alpha+\beta), \qquad \mathbb{E}[\log(1-p)] = \psi(\beta) - \psi(\alpha+\beta),$$

hence

$$\mathbb{E}\left[ \log \frac{p}{1-p} \right] = \psi(\alpha) - \psi(\beta).$$

Therefore,

$$\log \delta_n \xrightarrow{\text{a.s.}} (1-\gamma)\big[\psi(\alpha) - \psi(\beta)\big],$$

and exponentiating yields

$$\delta_n \xrightarrow{\text{a.s.}} \exp\big\{(1-\gamma)\big[\psi(\alpha) - \psi(\beta)\big]\big\} = \delta_\infty,$$

which completes the proof. □

### Acknowledgments

This research was supported by the German Research Foundation (DFG) under grants HA3000/29-1 and HA3000/30-1. I am grateful to Ulrike Hahn and Borut Trpin for helpful discussions, as well as the anonymous reviewers for their valuable feedback.

## References


Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*, 463–485.

Dietrich, F., & List, C. (2014). Probabilistic opinion pooling. In A. Hájek & C. Hitchcock (Eds.), *The oxford handbook of philosophy and probability* (pp. 403–429). Oxford University Press.

Eva, B., & Hartmann, S. (2018). Bayesian argumentation and the value of logical validity. *Psychological Review*, *125*(5), 806–821. https://doi.org/10.1037/rev0000101

Eva, B., Hartmann, S., & Rafiee Rad, S. (2020). Learning from conditionals. *Mind*, *129*(514), 461–508. https://doi.org/10.1093/mind/fzz010

Fennell, J., & Baddeley, R. (2012). Uncertainty plus prior equals rational bias: An intuitive bayesian probability weighting function. *Psychological Review*, *119*(4), 878–887. https://doi.org/10.1037/a0029346

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*(1), 129–166.

Herzog, S. M., & Hertwig, R. (2014). Harnessing the wisdom of the inner crowd. *Trends in Cognitive Sciences*, *18*(10), 504–506.

Juechems, K., Balaguer, J., Spitzer, B., & Summerfield, C. (2021). Optimal utility and probability functions for agents with finite computational precision. *Proceedings of the National Academy of Sciences*, *118*(2), e2002232118. https://doi.org/10.1073/pnas.2002232118

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

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

Pettigrew, R., & Weisberg, J. (forthcoming). Geometric pooling: A user's guide. *British Journal for the Philosophy of Science*. https://doi.org/10.1086/727000

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*, 864–901.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Vul, E., & Pashler, H. (2014). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *25*, 3–13.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Zhang, H., & Maloney, L. T. (2012). Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, *6*, 1. https://doi.org/10.3389/fnins.2012.00001

Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, *117*(36), 22024–22034. https://doi.org/10.1073/pnas.1922401117

Zhu, J.-Q., Sanborn, A. N., & Chater, N. (2020). The Bayesian sampler: Generic Bayesian inference causes incoherence in human probability judgments. *Psychological Review*, *127*(5), 719–748. https://doi.org/10.1037/rev0000190

Zilker, V., & Pachur, T. (2022). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*, *129*(5), 949–975. https://doi.org/10.1037/rev0000304