A Cognitive Theory of Reasoning and Choice

Pedro Bordalo, Nicola Gennaioli, Giacomo Lanzani, Andrei Shleifer*

April 29, 2026

Abstract

We present a theory of choice in which attention to the features of options is determined by the decision maker's categorization of the current problem in a set of problems she solved in the past. Categorization depends on goal-relevant and contextual problem-level features. The model yields heterogeneity in attention and choice in a given problem based on different past experiences and instability when changes in irrelevant context cause re-categorization. We show that heterogeneous and unstable representations of a choice problem unify major biases in judgment and decision making. *JEL codes:* D83, D91

*The authors are from Oxford Said Business School, Universita Bocconi, UC Berkeley, and Harvard University, respectively. We dedicate this paper to the memory of Danny Kahneman. We thank Daniel Bartels, John Conlon, Ben Enke, Thomas Graeber, Cosmin Ilut, Alex Imas, Lawrence Jin, Spencer Kwon, Shengwu Li, Jonathan Libgober, Ulrike Malmendier, Sendhil Mullainathan, Ryan Oprea, Andrea Prat, Joshua Schwartzstein, Jesse Shapiro, Philipp Strack, Tomasz Strzalecki, Dmitry Taubinsky, Richard Thaler, David Thesmar, Daniel Treisman, Robert Waldmann, Sevgi Yuksel, and four referees of this journal for extremely helpful comments. Steven Ma provided outstanding research assistance. Gennaioli thanks the European Commission for financial support through the ERC Advanced Grant (GA 101097578). Views and opinions expressed are those of the authors and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

1 Introduction

Is Donald Trump a hardened criminal or the champion of ordinary Americans? Are stocks an opportunity or a gamble? Is the car rental agency Avis a loser to Hertz or an underdog worth trying? Do guns at home make you safer from potential intruders, or less safe due to risk of an accident (Alsan, Schwartzstein, and Stantcheva [2])?

Economists trace choice to tastes or information. But even when these are constant, changing attention to features can alter decisions. Avis's advertising campaign "We are number two, we try harder" convinced many people to switch from Hertz to Avis not with new facts, but by changing what they think about when renting a car. What shapes how we think about a choice problem? How does this, in turn, affect the choice we make?

We present a theory addressing these questions. Our broad idea is that before solving a problem, a decision maker (DM) must ask: what kind of a problem is it? What matters for solving it? This process is fundamentally cognitive. We formalize this process and its impact on choice. In the problem recognition step, the DM tunes attention to features of the choice context (location, time, etc.) so as to maximize the similarity between her current problem and a category of problems she solved in the past. The selected category pins down attention to the attributes of choice options (price, quality, etc.), which forms her mental representation of the problem. In the second step, the DM chooses by maximizing her objectives given the mental representation. By shaping how the DM thinks about the problem, irrelevant context and past experiences create systematic heterogeneity and instability in choice, holding tastes and information fixed.

When asked to identify the letter in the middle of a word, people recognize "A" in C/-\T and "H" in T/-\E (Rumelhart and McClelland [105]), even though the task and the target stimulus are the same. Our theory explains this instability. In the first step, the adjacent letters prompt the DM to recognize the problem as solving "which word is this?". In turn, this category of problems focuses the DM on features of the string of letters, including the visual features of "/-\", that allow her to fit a known word, while neglecting discrepant features that may point to a different representation. Context and experiences point to a known letter (A or H), while the ambiguous symbol "/-\" in isolation may bring to mind a ladder or a street sign, rather than a letter.

The same process is at play in consumer choice. A jam offered in a standard jar in a supermarket aisle prompts recognition of "buying" a staple good, focusing attention on price. The same jam offered in a fancy jar on a wooden tasting board instead prompts recognition of "consuming" pleasure, focusing attention on quality. The jam problem could be recognized either way, depending on the location and other features, shaping attention to p or q. Three results follow.

First, changes in context cause sharp changes in valuation. A consumer recognizing choice as "buying" is insensitive to the jam quality relative to its price. By cueing a cozy breakfast experience at home, the fancy jar on the wooden board increases contextual similarity to "consuming", inviting a switch to that category. Attention to quality then increases relative to price, raising the jam's valuation. Price sensitivity is context-dependent because re-categorization causes a discontinuous fall in attention to p relative to q.

Second, a frequently used category is more likely to be used today, because the DM *endogenously* focuses on the context congruent with it. A formerly poor consumer is likely to notice that the jam costs money because this feature fits her frequent poverty experiences. Reliance on "buying" then causes her to be more price elastic than a consumer without the same experiences, who relies more on "consuming" and focuses on pleasure. A "price-focused" mental set can cause the neglect of large benefits, as when modest medical co-pays reduce demand for valuable care (Baicker, Mullainathan, and Schwartzstein [7], Chandra, Flack, and Obermeyer [34]). Experiences create heterogeneous price elasticities at constant income and tastes.

Third, a feature that is bottom-up salient, due to its prominence (Bushong, Camerer, and Rangel [29]) or contrast (Bordalo, Gennaioli, and Shleifer [25]), favors a switch to a category in which it is focal, reducing attention to other features. A consumer may neglect the risk that the jam is spoiled in normal times. Yet, a public concern that a few jars are contaminated makes quality salient, favoring a switch to "consuming". Attention is drawn to quality, including sickness from contamination, not to probabilities or price. As a result, valuation of jam falls dramatically and becomes insensitive to price. Re-categorization explains how unlikely risks can be either neglected or exaggerated, and leads to predictions for when each occurs.

In our theory, decision weights traditionally viewed as stable preferences (e.g., price sensitivity) or biases (e.g., probability weighting) are shaped by *how the DM recognizes and represents* the problem, which shapes attention to the attributes of a choice option. Section 2 introduces the model and show the unifying power of this logic. In Section 3 it generates heterogeneous and unstable price elasticities (Wakefield and Inman [126]) and mental accounting (Thaler [119]-[120]). In Section 4 it accounts for judgment biases typically explained by heuristics and their instability (Tversky and Kahneman [123], Bordalo, Conlon, Gennaioli, Kwon, and

Shleifer [20], Ba, Bohren, and Imas [6]), as well as for narratives (Andre, Roth, Haaland, and Wohlfart [3] Andre, Pizzinelli, Roth, and Wohlfart [4]). In Section 5 the logic yields, through bottom-up salience, a theory of "framing effects" (Tversky and Kahneman [124]) in both riskless choice (Enke and Zimmermann [45]) and risky choice (Kahneman and Tversky [74]).

By modelling the problem recognition *process*, we link these biases to cognitive proxies such as contextual similarity and experiences. The hallmark of categorization is within person multi-modal attention shaped by context: a person's attention to price compared to quality and thus her price elasticity should fall when context becomes more similar to consuming, for instance due to an attractive shopping environment. The hallmark of experiences is reliance on familiar categories: people who experienced painful losses are likely to focus on potential losses in contextually similar problems, exhibiting risk aversion. Cognitive proxies allow a researcher to separate representations from tastes and information.

Recent work studies bottom-up salience driven by prominence (Li and Camerer [81]) and contrast (Bordalo, Gennaioli, and Shleifer [21], Kőszegi and Szeidl [77]), and top-down attention shaped by goal optimality (Sims [116], Gabaix [52]) or priors (Schwartzstein [110], Gagnon-Bartsch, Rabin, and Schwartzstein [54], Schwartzstein and Sunderam [111]). Related approaches study insensitivity due to noise or uncertainty (Woodford [127], Enke and Graeber [44]). Our innovation is to show how recognition affects attention top down, based on payoff-irrelevant context and experiences, and how recognition delivers framing effects through the interaction of top-down attention with bottom-up salient stimuli.

Most closely related to this paper is Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20], who document multimodality and instability in choice and attention in statistical problems. For instance, in the balls-and-urns problem (Edwards [41]), people do not integrate the base rate and the likelihood into a potentially distorted Bayesian judgment. They bimodally cluster either on the base rate or on the likelihood, with measured attention also clustering on these features. The authors propose a model of probability judgment in which attention is driven by bottom-up contrast and "prominence." The current paper generalizes this model to non statistical problems and micro-founds prominence through recognition.

We build on the psychology of similarity perceptions (Nosofsky [97], Tversky [121]-[122]), top-down attention (Itti and Baldi [69], Awh, Belopolsky, and Theeuwes [5]), and categories (Mack and Palmeri [88], Reed [101], Rosch and Lloyd [103]). In Gilboa and Schmeidler [56], Mullainathan [94], and Fryer and Jackson [49], the DM pays full attention and uses knowledge from categories or similar problems to resolve uncertainty over an unknown attributes of current options. In our model, in contrast, there is no uncertainty and attention endogenously changes with experiences and salient stimuli, leading to choice heterogeneity and instability even in simple problems with full data. The resulting distortions then take the form of dampened or amplified sensitivity to a known feature across options, unifying base rate neglect, distorted price elasticities for known goods, and framing. Laibson [79] models consumption habits as context driven changes in reference points. In Samuelson [107], a DM co-categorizes dictator and bargaining games. Neither offers a general theory of problem recognition.[1]

Recent work on memory studies selective retrieval of the features of choice options (Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [19], Fudenberg, Lanzani, and Strack [50]-[51], Wachter and Kahana [125], Bordalo, Gennaioli, and Shleifer [24]). In our model feature values are known, and memory shapes attention they receive. Future work may study the two mechanisms together.

2 The Model

Section 2.1 describes how attention shapes valuation and choice. Section 2.2 describes how categorization and attention are jointly determined by recognition.

2.1 Attention and Valuation

A problem P includes: i) a nonempty finite menu O of lotteries defined on a finite probability space $(\Omega, 2^{\Omega}, \mathbb{P})$ and ii) a context vector $\kappa_P = (\kappa_{P,i})_{i \in \Phi_K}$ that reports features common to all lotteries and belonging to a finite set Φ_K.

Each lottery $o \in O$ is a finite set of vectors $y = (u, e)$, called atoms. Subvector u reports the numerical values of a finite set Φ_H of payoff carriers (e.g., quality, price) that we call *hedonic features*, so $u \in \mathbb{R}^{\Phi_H}$. Subvetcor e reports the values of a finite set Φ_E of *event features*, where each $i \in \Phi_E$ is a partition of Ω (e.g., selected urn, color of drawn ball). Thus $e = (e_i)_{i \in \Phi_E}$ reports the cell in each partition in which u is delivered.[2]

[1] In Salant and Rubinstein [106] instability is due to the use of different choice functions for different frames. Ellis and Masatlioglou [42] axiomatize utility that is stable within but not across categories of alternatives. Important work on categorization includes Mohlin [93] and Jehiel [70].

[2] Formally, in atom $y = (u, e)$ let $e_i \subseteq \Omega$ be a generic entry of e, $i \in \Phi_E$. Payoff u is delivered in event $\cap_{i \in \Phi_E} e_i$, whose probability is $\mathbb{P}(e) = \mathbb{P}(\cap_{i \in \Phi_E} e_i)$.

To illustrate, consider a lottery delivering a jam and a monetary payoff contingent on the outcomes g or b of a wheel of fortune W. In state g, the lottery delivers a good jam giving q_g utils and a monetary payoff m_g giving ηm_g utils, where $\eta > 0$ is the marginal utility of money. In state b it delivers a spoiled jam giving $q_b \leq q_g$ utils and a monetary payoff m_b giving ηm_b utils. The money payoff can be the jam's price, $m_g = m_b = -p < 0$, but the example covers monetary lotteries if m_g and m_b are general and $q_g = q_b = 0$.

There are four hedonic features: the good and bad jam qualities Qg and Qb, the corresponding money payoffs Mg and Mb. Each feature takes value 0 in the state where it is not delivered (e.g., Qg takes value 0 in state b and q_g in g, etc.). The event feature is the wheel outcome, $e \in \{g, b\}$. Thus, atom (u, e) reports in $u \in \mathbb{R}^4$ the value of hedonic features in state e. There are two atoms: one for g, one for b.[3]

In expected utility theory (EU), this description pins down valuation: the DM multiplies the total hedonics of (u, e) by the probability of e and adds across events. In our model, valuation depends also on the representation of P: a vector of attention weights $\alpha_P \in [0,1]^{\Phi}$, where $\Phi = \Phi_H \cup \Phi_E \cup \Phi_K$ is the set of hedonic, event, and context features. Context weights $\alpha_{P,i \in \Phi_K}$ shape problem recognition. We formalize them later.

The other weights shape the DM's sensitivity to her tastes u and to information about e. The perceived value of hedonic feature $i \in \Phi_H$ is, given weight $\alpha_{P,i}$, equal to:

$$u_i(\alpha_P) = \alpha_{P,i} \cdot u_i + (1 - \alpha_{P,i}) \cdot \overline{u}_i, \tag{1}$$

where u_i is the value of the hedonic in the atom and $\overline{u}_i$ is its average value across atoms in O.[4] For instance, less than full attention to the jam's price, $\alpha_{P,Mj} < 1$, $j \in \{g, b\}$, reduces perceived price variation in the choice set.

With a single event feature W, the perceived probability of its realization e is, given weight $\alpha_{P,W}$, equal to:

$$\Pr(e; \alpha_P) = \frac{\mathbb{P}(e)^{\alpha_{P,W}}}{\sum_{e' \in W} \mathbb{P}(e')^{\alpha_{P,W}}}. \tag{2}$$

Less than full attention, $\alpha_{P,W} < 1$, causes the DM to smear the probability distribution of W towards a uniform, fully so if $\alpha_{P,W} = 0$. The DM overestimates the event that the jam is spoiled if she does not attend to its small numerical probability.

[3] The atoms are $(q_g, \eta m_g, 0, 0, g)$ and $(0, 0, q_b, \eta m_b, b)$. The spoiled jam is not delivered in state g so its value in this state is zero, and similarly for other non-delivered features.

[4] Formally, $\overline{u}_i = \frac{\sum_{o \in O} \sum_{(u,e) \in o} u_i}{\sum_{o \in O} |o|}$.

The value of atom y under representation α_P is the product of its perceived probability in Equation (2) and its total perceived payoff computed using Equation (1):

$$v(y;\alpha_P) = \Pr(e;\alpha_P) \sum_{i\in\Phi_H} u_i(\alpha_P), \tag{3}$$

and the value of lottery o is the sum of its atomic values $v(o;\alpha_P) = \sum_{y\in o} v(y;\alpha_P)$.

In our example, up to a choice-irrelevant affine transformation, valuation satisfies:

$$v(o;\alpha_P) = \sum_{e\in\{g,b\}} \mathbb{P}(e)^{\alpha_{P,W}} \left[\alpha_{P,Qe}(q_e - \overline{q}_e) + \alpha_{P,Me}\eta(m_e - \overline{m}_e)\right]. \tag{4}$$

If attention is full, $\alpha_P = \mathbf{1}$, we recover expected utility. If attention is not full, Equation (4) nests well-known behavioral models.

If the jam is always good ($\mathbb{P}(b) = 0$) and if $m_g = -p$, we have riskless choice with weighted utility, $v(o;\alpha_P) = \alpha_{P,Q}(q - \overline{q}) - \alpha_{P,M}\eta(p - \overline{p})$. Attention to money relative to quality $\alpha_{P,M}/\alpha_{P,Q}$ pins down price sensitivity, as in Bordalo, Gennaioli, and Shleifer [22]. Critically, if the DM neglects the spoiled jam payoffs ($\alpha_{P,Qb} = \alpha_{P,Mb} = 0$), she chooses "as if" choice is riskless even if $\mathbb{P}(b) > 0$.

In monetary lotteries ($q_e = 0$ for all $e \in \{g,b\}$), inattention to the DGP, $\alpha_{P,W} < 1$, gives overweighting of unlikely events (Kahneman and Tversky [74]). Higher attention to the upwside vs downside monetary payoff $\alpha_{P,Mg}/\alpha_{P,Mb}$ yields payoff-driven risk seeking (Bordalo, Gennaioli, and Shleifer [21]).

This setting also embeds statistical problems: a hypothesis is a lottery that delivers a payoff of \$1 if an outcome of W occurs and zero otherwise. We later allow for multi-dimensional e, which produces long-standing judgment biases.

Crucially, attention weights to features of tastes and information, $(\alpha_{P,i})_{i\in\Phi_H\cup\Phi_E}$, depend on attention to context, weights $(\alpha_{P,i})_{i\in\Phi_K}$. The reason is that attention to context shapes problem recognition, as we show next.

2.2 Recognition and Choice

The context vector κ_P reports standard relevant features of P, such as the choice set, the range of hedonics (e.g., available qualities or prices), and the list of events.[5] It also reports irrelevant features such as time, location, or packaging, which can affect recognition. If the jam is offered with rich sensory features (e.g., a fancy jar)

[5] Features in κ_P can also include the average price level (expensive versus cheap goods problem) or the average probabilities of specific events (high versus low risk problem).

and the DM attends to them, she represents the problem P as "consuming" rather than "buying". This prompts focus on jam qualities and neglect of prices or risk. If the DM attends to the supermarket location, she recognizes P as "buying", causing focus on price and neglect of quality.

Attention to context and recognition depend on the categories of problems the DM solved in the past. To see how, we summarize the current problem as a vector of attention plus context (α_P, κ_P).

Categories. The DM's memory database is partitioned into a set of categories C. Past problems in category $c \in C$ are also summarized by an attention plus context "prototype" vector (α_c, κ_c) and by their time-discounted frequency $F_c \in \mathbb{R}_+$.[6]

Attention α_c is the prototypical representation used in category c problems. It specifies the context diagnostic of experiences in c (Rosch and Lloyd [103]) via binary attention $\alpha_{c,i} \in \{0,1\}$ for $i \in \Phi_K$. Eating jam mostly occurs at "home", but in a wide range of weather conditions, so location feature i is diagnostic for this category, $\alpha_{c,i} = 1$, but weather feature i' is not, $\alpha_{c,i'} = 0$. Vector α_c also reports the attention $\alpha_{c,i}$ the DM paid to hedonics and events in c, $i \in \Phi_H \cup \Phi_E$. When eating jam, the DM is more focused on quality and less on price than when buying jam. Experiences "glue" context with attention to relevant features.

Recognition. To solve P, the DM must decide "what kind of problem is this? What matters for solving it?" To this end, she chooses attention α_P and a category c to maximize the total similarity between the current problem (α_P, κ_P) and a category prototype (α_c, κ_c). Similarity is measured by a function $S\left[(\alpha_P, \kappa_P), (\alpha_c, \kappa_c)\right]$ that decreases if: i) the context κ_c of past problems in c is different from current context κ_P, or ii) the representation α_c used for problems in c differs from the current one α_P. We use the separable form:

$$S\left[(\alpha_P, \kappa_P), (\alpha_c, \kappa_c)\right] = 1 - \frac{\sum_{i\in\Phi} d\left(|\alpha_{P,i} - \alpha_{c,i}|\right) + \sum_{i\in\Phi_K} \alpha_{P,i}\alpha_{c,i} d_i\left(\kappa_{P,i}, \kappa_{c,i}\right)}{|\Phi| + |\Phi_K|}. \quad (5)$$

Discrepancies in attention are measured by $d : [0,1] \to [0,1]$, which is strictly increasing and convex. Discrepancies in context feature $i \in \Phi_K$ are measured by a distance d_i with values in $[0,1]$. This functional form helps tractability, but our results rely on the general idea – central in psychology (Nosofsky [97]) – that similarity

[6]For a DM facing a problem in period $t \in \mathbb{N}$, a category $c \subseteq \{1, ..., t-1\}$ collects a set of past periods. Recency-weighted frequency is $F_c = \sum_{\tau\in c} \delta^{t-\tau}$, $\delta \in (0,1)$. The "prototype" can be interpreted as the average attention $\alpha_c = \sum_{\tau\in c} \alpha_\tau / |c|$ used in these problems and the "average" context κ_c where, for every $i \in \Phi_K$, $\kappa_{c,i}$ minimizes some discrepancy from $(\kappa_{\tau,i})_{\tau\in c}$.

falls in differences, and more so when these are more attended to.[7]

Total similarity multiplies S by F_c, capturing the idea that more familiar categories, higher F_c, are easier to retrieve. Similarity is also perturbed by a type I extreme-value shock ϵ_c with precision parameter $\lambda > 0$, reflecting random attention to categories. This error structure is not necessary but yields convenient closed forms.

Choice. Conditional on a draw $(\epsilon_c)_{c\in C}$, the DM solves P as follows:

$$\left(\alpha_P^*, c^*\right) = \text{argmax}_{\alpha_P\in[0,1]^\Phi, c\in C} \; F_c S\left[\left(\alpha_P, \kappa_P\right), \left(\alpha_c, \kappa_c\right)\right] + \epsilon_c, \tag{6}$$

$$o^* = \text{argmax}_{o\in O} \; v\left(o; \alpha_P^*\right). \tag{7}$$

In Equation (6) the DM represents the problem using category c^*. She then uses the endogenous weights α_P^* to value and pick the best lottery o^* in Equation (7). In rational choice, attention is full $\alpha_P^* = \mathbf{1}$ and the DM maximizes $v\left(o; \mathbf{1}\right)$. In our model, keeping $v\left(o; \mathbf{1}\right)$ fixed, attention and choice are heterogeneous due to experiences F_c, and unstable due to context κ_P.

Rational accounts of heterogeneity and instability assume that hedonics in $v\left(o; \mathbf{1}\right)$ directly depend on experiences through habits, as in Stigler and Becker [117], or on context such as advertising, as in Becker and Murphy [14]. This account is tautologically consistent with choice because tastes are not measured directly, only choices are. A deeper understanding of choice can be achieved by measuring beliefs, the probabilities in $v\left(o; \mathbf{1}\right)$, for instance about returns to human capital (Manski [92]) or financial assets (Bordalo, Gennaioli, La Porta, and Shleifer [26, 27]). Measured beliefs narrow down unobserved variation in tastes. Our model highlights a deeper force: shifting attention to features. The way the DM thinks about a problem can affect her weighting of hedonics and events, which may be confused with differences in tastes and information, respectively. Structure on α_P is needed to identify the drivers of choice.

Equation (6) imposes such a structure and links the theory's endogenous objects, attention and choice, to proxies for the cognitive process: experiences, features of context, and similarity. A growing body of work measures these proxies in the field (for instance, using administrative and text data), in the lab, or in surveys, and shows that – even if they are unrelated to choice-relevant tastes or information – they drive behavior through selective attention or memory mechanisms (Taubinsky, Butera, Saccarola, and Lian [118], Cenzon [33], Fan, Liang, and Peng [46], Ba,

[7]Consistent with an attention-dependent metric, Goldstone [58] shows that people trained to relate specific features among objects in one task subsequently perceive differences across those features to be larger, and differences across other features to be smaller.

Bohren, and Imas [6]). For consumer choice, Bordalo, Burro, Gennaioli, Nacamulli, and Shleifer ([18]) show that advertising of a brand increases consumption of other brands with similar ads, especially for consumers who have more experience with those brands. In belief formation, Bordalo, Gennaioli, Lopez de Silanes, Schroder, Shleifer, and van Rooij [28] show that macroeconomic narratives and the entailed expectations are influenced by recall of uninformative personal experiences, especially in contextually similar domains.

Our model shows how to use these proxies to study selective "top-down" attention driven by problem recognition, which shapes choice even if domain relevant tastes or information are fully available to the DM, such as when choosing at a supermarket between two known brands or when evaluating two statistical hypotheses based on a described DGP. The problem recognition step makes further predictions on the reasoning and narratives used to justify choices, which are now routinely measured in surveys (Haaland, Roth, Stantcheva, and Wohlfart [62], Andre, Haaland, Roth, and Wohlfart [3]), and on how these depend on bottom up salient features (Li and Camerer [81]).

We next illustrate our model's general predictions on representations. We then link these predictions to choices and judgments.

2.3 General Predictions on Representations

Consider how the DM sets optimal attention α_P to match problem P to a given category c. By Equation (6) an interior solution for a generic feature $i \in \Phi_K$ satisfies the first order condition:

$$d'\left(\alpha_{c,i} - \alpha_{P,i}\right) = d_i\left(\kappa_{P,i}, \kappa_{c,i}\right)\alpha_{c,i}. \tag{8}$$

Setting a representation α_P close to the categorical one α_c increases similarity to c (the left hand side term), but may highlight differences in context diagnostic of the category, $d_i\left(\kappa_{P,i}, \kappa_{c,i}\right)\alpha_{c,i} > 0$ on the right hand side. We then obtain:

Proposition 1 *Optimal attention at fixed category c,*

$$\alpha_P(c) = \text{argmax}_{\alpha_P \in [0,1]^\Phi} S\left[\left(\alpha_P, \kappa_P\right), \left(\alpha_c, \kappa_c\right)\right],$$

is equal to category attention, $\alpha_{P,i}(c) = \alpha_{c,i}$, *unless feature* $i \in \Phi_K$ *is diagnostic of c and discrepant with P,* $d_i\left(\kappa_{P,i}, \kappa_{c,i}\right) \cdot \alpha_{c,i} > 0$. *In the latter case, attention to* *that feature is shrunk toward zero, $\alpha_{P,i}(c) \leq \alpha_{c,i}$, and the more so the higher is $d_i(\kappa_{P,i}, \kappa_{c,i})$.*[8]

For hedonics and events, attention fully adapts to the category's α_c. The choice implication is that the sensitivity of the DM's valuation to price, qualities, and so on, is shaped by the category.

Attention is instead lower than in α_c for contextual features where the problem differs from the category, the more so the greater the difference. As a consequence, the perceived similarity of problem P to a category c is endogenous, given by $S(P,c) = F_c[1 - d(P,c)]$, where

$$d(P,c) = \min_{\alpha_P \in [0,1]^\Phi} \frac{\sum_{i\in\Phi} d(|\alpha_{c,i} - \alpha_{P,i}|) + \sum_{i\in\Phi_K} \alpha_{P,i}\alpha_{c,i} d_i(\kappa_{P,i}, \kappa_{c,i})}{|\Phi| + |\Phi_K|}$$

is the minimized distance from c. Endogenous inattention reduces the sensitivity of similarity $S(P,c)$ to contextual discrepancies $d_i(\kappa_{P,i}, \kappa_{c,i})$, with important implications. Equilibrium representations work as follows.

Proposition 2 *Suppose that $\alpha_P(c) \neq \alpha_P(c')$ for all $c \neq c'$. The DM's representation satisfies*

$$\Pr(\alpha_P^* = \alpha_P(c)) = \frac{\exp\{\lambda F_c[1 - d(P,c)]\}}{\sum_{c'\in C} \exp\{\lambda F_{c'}[1 - d(P,c')]\}} \qquad \forall c \in C. \tag{9}$$

Problem P is more likely to be represented using $\alpha_P(c)$ when category c:
i) has a context closer to the current one, $\partial \Pr(c|P)/\partial d_i(\kappa_{P,i}, \kappa_{c,i}) \leq 0$.
ii) was used more frequently or recently, $\partial \Pr(c|P)/\partial F_c \geq 0$.

Due to the shock ϵ_c categorization is stochastic, causing discrete switches in attention to options' attributes. Our first behavioral prediction follows: within person multi-modal valuation. In similar inference problems, a person can switch from using only the base rate to using only the likelihood, as shown in Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20]. This pattern does not arise with noisy perception of payoffs/probabilities (Woodford [127]) or of the optimal action (Enke and Graeber [44]). In these models, the weighting of the features is stable, so valuation is uni-modal. In the data it is not, and our model explains why.

[8] Here optimal attention is unique. The appendix relaxes this assumption.

Relatedly, context creates systematic instability: a person is less likely to use a category c when a current context feature $\kappa_{P,i}$ becomes more discrepant from the category's $\kappa_{c,i}$. A second behavioral prediction follows: valuation predictably switches from one mode to another after changes in context. Changing the inference context from a balls an urns problem (Edwards [41]) to a normatively identical taxicabs problem (Kahneman and Tversky [73]) causes many people to switch from the base rate to the likelihood even though statistics are unchanged (Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20]).

Finally, experiences create systematic heterogeneity: people more familiar with a category c (higher F_c) are more likely to use it. A third behavioral prediction follows: valuation differs across people, due not just to tastes but also to differential attention to features. In inference problems, some people are more likely to use the base rate or the likelihood than others despite common incentives and information (Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20]). This heterogeneity reflects the fact that people endogenously neglect different features of context (those that are dissonant with the categories they are familiar with). It also implies that the overall distribution of valuations may appear unimodal (if many categories are possible) even if individual distributions are not.

Our theory generates famous problem solving errors (Bassok [12]). First, training on a problem improves performance but expertise is *too context specific* (Camerer and Johnson [30], Chase and Simon [36], Green, Rao, and Rothschild [60]). A DM trained on inference with numerical priors and likelihoods can reach the correct solution with arbitrary values and yet fail to apply Bayes rules or conditioning when the problem – and hence its context – has no numbers (such as in Kahneman and Tversky's Linda problem [75]). Second, experience with a problem (high F_c) causes *overgeneralization*: the familiarity of some people with solving inference problems may cause them to focus on the empirical share of good to bad signals even when this statistic is irrelevant, as in the case of the Gambler's Fallacy studied in Section 4.

We next derive the implications of our model for riskless choice, statistical problems, and risky choice. Specific predictions depend on the set of categories C, which we take as given, as does Mullainathan [94]. To discipline our analysis we restrict attention to four categories, two in riskless choice, two in statistical problems. The conclusions discuss ways to endogenize and measure a broader set of categories.

3 Representations in Riskless Choice

The DM must choose whether or not to buy a jam. The state is not yet realized, so we are in the case of Equation (4). The modal state g exhibits "normal" quality $q_g = q$ and price $p_g = p$. State b is an unlikely "shock" that the DM has never experienced, and that throughout this section will be assumed to have subjective probability $\mathbb{P}(b) = 0$, so that its quality q_b and price p_b do not play a role.[9] Context is a unidimensional situation, feature K, with value $\kappa_P \in [0, 1]$. There are two categories: "buying" and "consuming."

Buying. This category, denoted *buy*, reflects experiences in which the DM focuses on the pain of paying for jam and partly on its forecasted quality (otherwise she would not buy), so $\alpha_{Mg} = 1 > \alpha_{Qg} = \alpha$. The unlikely shock b is neglected, $\alpha_{Qb} = \alpha_{Mb} = 0$. Diagnostic context embeds various features such as being at the supermarket, captured by $\kappa_{buy} = 0$. The *buy* prototype $(\alpha_{buy}, \kappa_{buy})$ is thus the vector $(\alpha_{buy,Qg}, \alpha_{buy,Mg}, \alpha_{buy,Qb}, \alpha_{buy,Mb}, \alpha_{buy,W}, \alpha_{buy,K}, \kappa_{buy}) = (\alpha, 1, 0, 0, \alpha_{buy,W}, 1, 0)$.

Consuming. This category, denoted *con*, reflects experiences in which the DM focuses on the pleasure of jam, while paying less attention to the purchase price. Ex-post, the DM focuses on the realized quality during consumption, ex-ante she focuses on the modal quality. From the ex-ante perspective, which is relevant for our current exercise, the DM attends to quality and partially to price, $\alpha_{Qg} = \alpha_{Qb} = 1 > \alpha = \alpha_{Mg} = \alpha_{Mb}$. Diagnostic context reflects features of tasting, such as being in one's kitchen, captured by $\kappa_{con} = 1$. The consuming prototype is then $(\alpha_{con,Qg}, \alpha_{con,Mg}, \alpha_{con,Qb}, \alpha_{con,Mb}, \alpha_{c,W}, \alpha_{con,K}, \kappa_{con}) = (1, \alpha, 1, \alpha, \alpha_{con,W}, 1, 1)$.

In this setup, the experiences in *con* entail an accurate perception of normal quality q, while experiences in *buy* entail an accurate perception of normal price p and choice is seen as riskless. To properly trade off q and p the DM should integrate the valuation of quality in *con* and of price in *buy*. However, *con* and *buy* are segregated in different contexts κ_c and they compete for the DM's representation based on: i) the DM's familiarity F_c and ii) the current context κ_P.

[9] This will be relaxed in Section 5.2.

3.1 The Context Dependent Consumer

With attention distance $d(x) = x/2$, and context distance $d_Z(z, z') = |z - z'|$, where Z denotes the single context dimension in κ_P, e.g., location, we have:

$$S[(\alpha_P, \kappa_P), (\alpha_c, \kappa_c)] = \quad (10)$$
$$1 - \frac{[|\alpha_{P,Qg} - \alpha_{c,Qg}| + |\alpha_{P,Mg} - \alpha_{c,Mg}| + |\alpha_{P,Z} - 1| + 2\alpha_{P,Z} \cdot |\kappa_P - \kappa_c|]}{8}.$$

For each category $c \in \{con, buy\}$, the DM sets the attention $\alpha_P(c)$ that maximizes similarity, producing two competing representations. The resulting similarities to P give, as a function of current context κ_P, the following result.

Proposition 3 (Representation and Similarity) *Given optimal weights $\alpha_P(c) = (\alpha_{Qg}(c), \alpha_{Mg}(c), \alpha_Z(c))$ on normal quality, price, and on context for category c, we have:*

1. *To represent P as buy, the DM sets $\alpha_P(buy) = (\alpha, 1, 1)$ if $\kappa_P < 1/2$ and $\alpha_P(buy) = (\alpha, 1, 0)$ if $\kappa_P > 1/2$. Thus, $S(P, buy) = F_{buy} \max[1 - \kappa_P/4, 7/8]$.*

2. *To represent P as con, the DM sets $\alpha_P(con) = (1, \alpha, 1)$ if $\kappa_P > 1/2$ and $\alpha_P(con) = (1, \alpha, 0)$ if $\kappa_P < 1/2$. Thus, $S(P, con) = F_{con} \max[(3 + \kappa_P)/4, 7/8]$.*

The DM focuses on the jam's price and neglects its quality when using *buy*, and the reverse with *con*. Moreover, she focuses on context when it resembles the category, namely low κ_P for *buy* and high for *con*, and neglects it otherwise. Neglecting the problem's context reduces similarity to the category (the DM *knows* that context should match), but focusing on the associated discrepancy reduces similarity even more.

Maximal similarity $S(P, c)$ increases in the category's frequency F_c and decreases in the distance between its context κ_c and κ_P. With negligible shocks ϵ_c, $\lambda \to \infty$, recognition in Equation (6) yields the following result.

Proposition 4 (Categorization and Valuation) *The DM categorizes P as buy if $\kappa_P < \kappa\left(\frac{F_{buy}}{F_{con}}\right)$ and as con otherwise, where $\kappa(\cdot)$ is increasing. Valuation satisfies:*

$$v(jam; \alpha_P^*) = \begin{cases} \alpha q - \eta p & \text{if} \quad \kappa_P < \kappa\left(\frac{F_{buy}}{F_{con}}\right) \\ q - \alpha\eta p & \text{if} \quad \kappa_P \geq \kappa\left(\frac{F_{buy}}{F_{con}}\right) \end{cases} . \quad (11)$$

The DM recognizes P as "buying" if context κ_P is sufficiently similar to $\kappa_{buy} = 0$, and as consuming otherwise. *buy* more likely prevails if this category is relatively more frequent, higher F_{buy}/F_{con}. Figure 1 assumes that the categories are equally frequent $F_{buy} = F_{con} = F$ and plots: i) the similarity of P to each category in Panel A and ii) the valuation function (11) in Panel B. Two properties stand out.

First, valuation of the jam is unstable. Categorization changes from *buy* to *con* as context crosses $\kappa_P = 1/2$ from below, causing attention to quality and valuation to shift upward. The DM is "stingy at the supermarket" ($\kappa_P < 1/2$) and may fail to buy even if $q > \eta p$, but "neglects opportunity cost at home" ($\kappa_P > 1/2$), enjoying the jam even if $q < \eta p$. Second, price sensitivity is high for $\kappa_P < 1/2$ and low for $\kappa_P > 1/2$. A vertically superior jam located on a rational indifference curve $q - \eta p = u$ is under-valued in *buy* and over-valued in *con*.

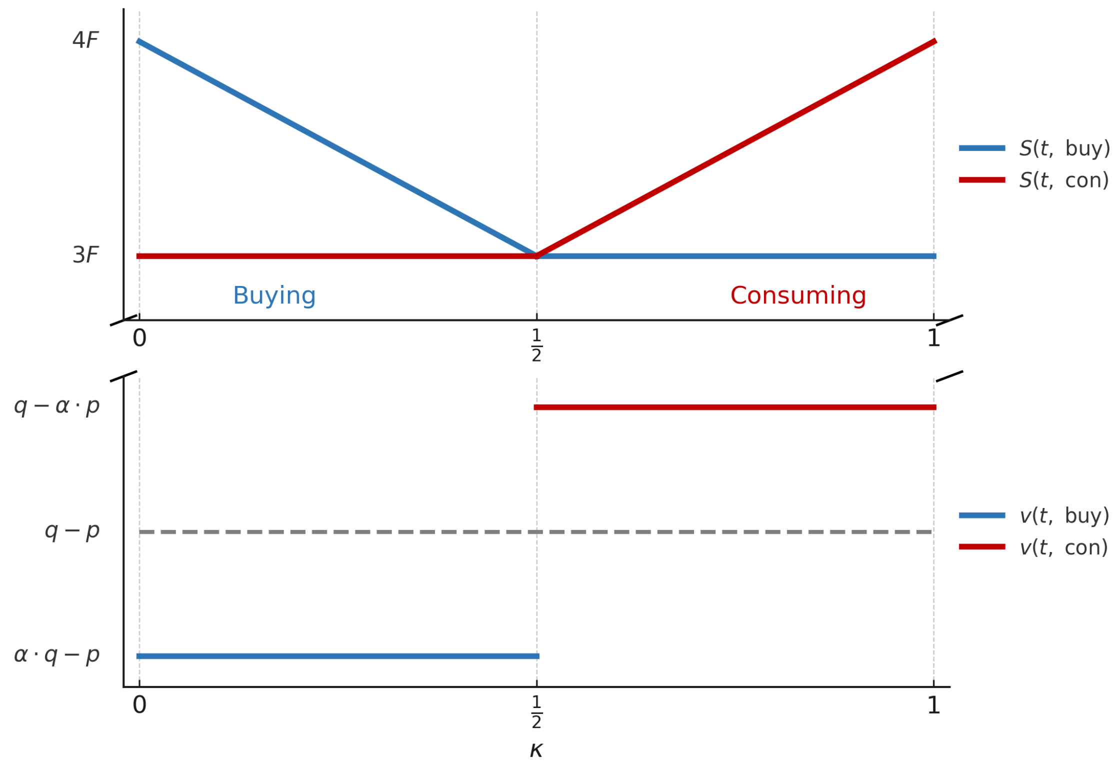
First model prediction: a person's price sensitivity should weakly decline if context becomes more similar to *consuming* experiences and less so to *buying* ones. Selling the jam in a fancy jar on a wooden tasting board evokes special occasions of enjoying treats at home, while reducing similarity to buying staple products, boosting sensitivity to quality and reducing that to price.

This predictions accounts for realistic instability in price elasticity. Wakefield and Inman [126] show that a person is price inelastic if she categorizes the good as "hedonic" and price elastic if she categorizes the good as "functional". This phenomenon reflects the role of context, embedded in the good itself or in the situations in which it is typically consumed. The context of a "friday night takeout" prompts focus on pleasure due to past treat experiences, encouraging people to order an extra appetizer or dessert. The context of a work lunch prompts focus on minimizing spending, increasing the price sensitivity.

Context based instability explains why uninformative advertising and spurious differentiation can tranform a good from "functional" to "hedonic": they increase the similarity of context to *con*, leading to the relative neglect of price. By creating a "living room" experience, a comfortable cafe draws attention to coffee taste and sociality, increasing valuation even for people who get the drink to go. By drawing attention to sources in tropical island or glaciers, branded water draws attention to its freshness, increasing valuation even for a standard product. A context driven shift from "buying" to "consuming" causes de-commoditization.[10]

[10] Bordalo, Gennaioli, and Shleifer [23] first explored de-commoditization, and its inverse commoditization, in terms of selective attention to quality and prices respectively. The mechanism there is bottom up contrast in a vertically differentiated market. In Section 5 we examine how bottom up drivers of salience complement the categorization mechanism.

Figure 1:

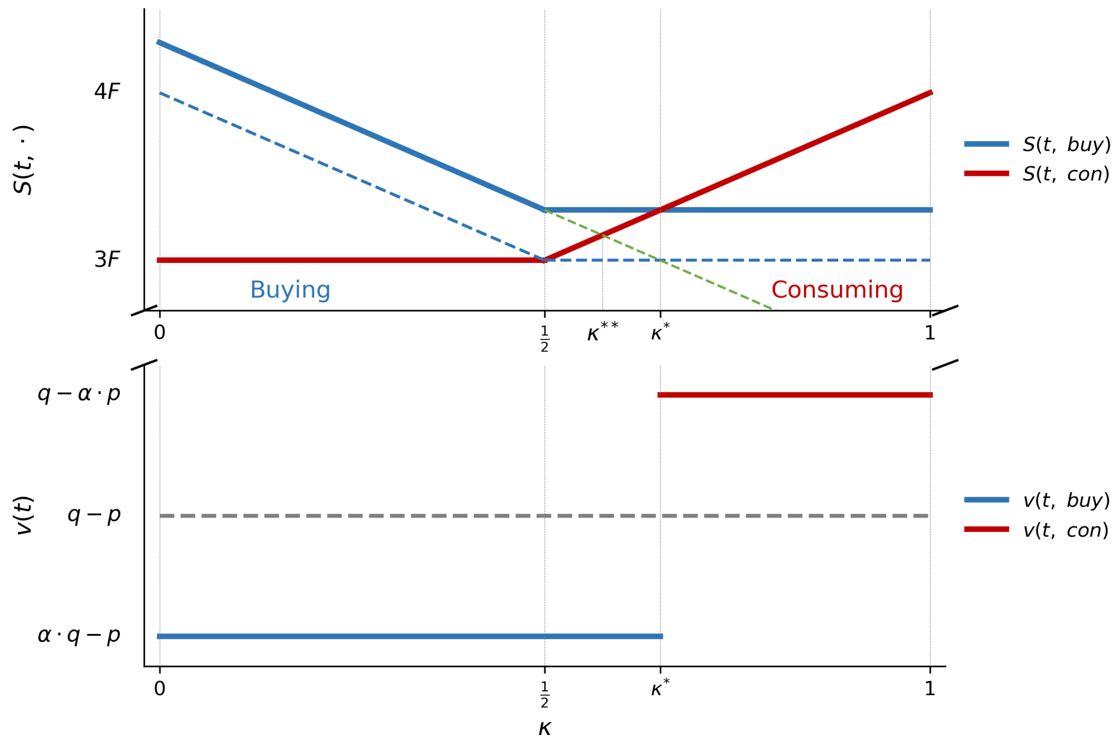


Panel A plots similarity $S(P, c)$ to categories *con* and *buy*, Equation 10. Panel B plots valuation, Equation (11), that follows from categorization into the most similar category.

The model also makes predictions based on experiences. Suppose that the frequency of *buy* is higher than that of *con*, $F_{buy} > F_{con}$. In Figure 2 the similarity to *buy* uniformly increases from the dashed to the solid blue line, promoting that category at any context κ_P. The consumer is more price-elastic even if her true marginal utility of money η is the same. Endogenous attention causes insensitivity to discrepancies with *buy*, leading to a much higher threshold κ^*, also relative to the cutoff κ^{**} attained if the DM fully attends to context (green line in Figure 2).

Second model prediction: Take two people A and B having the same income and tastes (same η and q). The person with the higher relative frequency of buy experiences (higher F_{buy}/F_{con}) should be weakly more price elastic in any context κ_P.

Figure 2:



Increasing the frequency F_{buy} increases similarity $S(P, buy)$, increasing the range of context κ_P where P is categorized as *buy*. This includes the range $\kappa_P \in \left(\frac{1}{2}, k^*\right)$ where the DM maximizes similarity by neglecting discrepant context.

Familiarity with reasoning about prices prompts the DM to recognize the current choice as "buying", increasing the price sensitivity.

The role of experiences explains the puzzling heterogeneity of price elasticities. People with experiences of poverty are more price elastic across situations (Shah, Zhao, Mullainathan, and Shafir [114]). Having experienced high opportunity costs, they tend to see price as highly relevant. This is overgeneralization, not a higher opportunity cost η or a Bayesian prior. It can cause large mistakes when quality q is high, such as saving on valuable medical treatment by focusing on low co-pays (Baicker, Mullainathan, Schwartzstein [7], Chandra, Flack, Obermeyer [34]). Price focus can persist even if the DM is no longer poor, contributing to variation in price elasticities and thriftiness for given income and wealth (Hoch, Kim, Montgomery, and Rossi [68], Rick, Cryder, and Loewenstein [102]). Measured experiences help test our mechanism: Link, Peichl, Pfauti, Roth, and Wohlfart [82] show that people with high inflation experiences attend more to inflation today.[11]

3.2 Mental Accounting

People use different accounts to track costs and benefits in different situations, leading to opportunity cost neglect, sunk cost fallacy, non-fungibility of money, and so on. Consider the following examples.

Opportunity Cost Neglect. Many years ago a person bought for \$20 a bottle of wine worth \$75 today. The person drinks the wine today. What is the cost she feels? Many people say zero or \$20 (Thaler [119]). They neglect the opportunity cost of drinking, the \$75 market price.

Sunk Cost Fallacy. A person bought a \$20 ticket to a football game that takes place a month later. On the game day, there is a severe blizzard. Does the person drive to the game? Would she drive had she been given the ticket for free? Frequent answers are "yes" to the first question, and "no" to the second, violating revealed preference: if the blizzard is severe enough, not driving to the game is optimal regardless of having paid the ticket price.

Opportunity cost neglect has been attributed to the temporal remoteness of the wine's purchase price (Gourville and Soman [59]), while the sunk cost fallacy has been explained with diminishing sensitivity (Thaler [119]) or distaste for "waste" (Shafir and Thaler [113]).[12]

We show that both phenomena arise from categorization in our baseline consumer problem, which yields new predictions. In both examples, we are in an ex-post situation where the "not-normal" state b has realized. In the wine problem, a shock affects the price, which increases by $\Delta p > 0$. In the football problem, a shock affects

[11] Context-based instability and experience-based heterogeneity interact in driving categorization. Instability can explain why even the poor can "splurge" in contexts associated with consumption pleasure, such as festivals or "treat" goods such as cigarettes (Banerjee and Duflo [8]), or why a person who grew up poor is less price elastic in new goods, such as i-Phones or travel, whose context is distant from her *buy* experiences in items such as clothes or food.

[12] Kőszegi and Matějka [76] offer a model of rational inattention which produces instances of behavior linked to mental accounting when hedonics are noisy. For substitute goods, attention is optimally focused on features that distinguish them, leading to category-budgets. The same mechanism leads to naive diversification if goods are complements. However, this model does not explain the sunk cost fallacy or the wine example, or more generally the role of context in driving instability.

quality, which changes by $\Delta q < 0$. Anomalies then arise because the consumer neglects these shocks due to categorization into consuming or buying.

In the wine example, the DM must evaluate the cost of drinking so both the original purchase price p and the capital gain Δp are relevant. In the football example, the DM must evaluate the benefit of driving to the game, so both the quality forecast q at the time of buying and the disutility of driving Δq are relevant. As a function of attention weights, valuations in the two problems are:

$$v(\text{cost of drinking}; \alpha_P) = \alpha_M p + \alpha_{\Delta M} \Delta p, \tag{12}$$
$$v(\text{driving}; \alpha_P) - v(\text{not driving}; \alpha_P) = \alpha_Q q + \alpha_{\Delta Q} \Delta q. \tag{13}$$

With full attention, decisions are rational. In the wine problem, $\alpha_M = \alpha_{\Delta M} = 1$ valuation equals the market price, $p + \Delta p = \$75$ (Equation 12). In the football problem, $\alpha_Q = \alpha_{\Delta Q} = 1$ the decision is to go to the game if and only if $q + \Delta q > 0$.

However, the consuming and buying categories do not attend to these features fully. *Con* fully attends to actual quality, which is now realized in b, so $\alpha_{con,Q} = \alpha_{con,\Delta Q} = 1$, but it only partially attends to the price at which the good was bought, $\alpha_{con,M} = \alpha \in (0,1)$, and it fully neglects the capital gain, which is not typically attended to when enjoying consumption, $\alpha_{con,\Delta M} = 0$. On the other hand, *buy* fully attends to the price paid and partially attends the quality forecast ex-ante, during the buying decision, $\alpha_{buy,M} = 1$ and $\alpha_{buy,Q} = \alpha \in (0,1)$. It however neglects both quality shocks and capital gains, $\alpha_{buy,\Delta Q} = \alpha_{buy,\Delta M} = 0$, which were not attended to when buying. These categories prompt evaluations:

$$v(\text{cost of drinking}; \alpha_P) = \begin{cases} \alpha p & \text{if } \alpha_P = \alpha_P(con) \\ p & \text{if } \alpha_P = \alpha_P(buy) \end{cases}, \tag{14}$$
$$v(\text{driving}; \alpha_P) - v(\text{not driving}; \alpha_P) = \begin{cases} (q + \Delta q) & \text{if } \alpha_P = \alpha_P(con) \\ \alpha q & \text{if } \alpha_P = \alpha_P(buy) \end{cases}.$$

In the wine problem, "consuming" focuses the DM on the pleasure of drinking, causing opportunity cost neglect. If $\alpha = 0$ the DM reports a zero cost, like many experimental subjects.[13] "Buying" instead focuses the DM on the purchase price p, which is also reported by several subjects. The capital gain is still neglected: it was not attended to when buying the bottle.

[13]During consumption, with prices not explicitly mentioned, we often feel no opportunity cost. Frederick, Novemsky, Wang, Dhar, and Nowlis [48] show that describing how savings can be used for other purchases substantially decreases the probability of spending.

In the football problem, "consuming" focuses the DM on the qualities of driving: seeing the game q but also on the blizzard $\Delta q < 0$, leading to the rational evaluation. "Buying" instead focuses the DM on the purchase price p and the normal benefit of seeing the game q. The blizzard shock is neglected: it was not attended to when buying the ticket.

The context κ_P of the vignette affects the modal answer by shaping similarity to different categories. In the wine problem, the description "drinks the wine" prompts similarity to *con* (high κ_P in Figure 1), favoring opportunity cost neglect. In the football problem, the choice of "going versus not going to the game" evokes a buying problem (low κ_P in Figure 1), promoting neglect of the blizzard, but less so if the ticket was given for free and did not involve a decision to buy.

This analysis then yields two comparative statics:

1. Frequency. People with higher F_{buy}, exhibit less opportunity cost neglect and more sunk cost fallacy. These could be people with poverty experiences or, in the wine example, people who have bought wine but have not yet drunk it (so recency weighted frequency of buying is high). In the football example, people who bought season tickets face one buying experience but many consuming ones; they have a high F_{con}, which undermines *buy* and reduces the sunk cost fallacy.

2. Instability. Making the blizzard more salient in the description or making an ex-ante plan for bad weather should increase reliance on *con*, reducing the sunk cost fallacy. Making the re-selling of the wine more salient, either as a forgone opportunity with consumption or as a possibility when buying, should hinder *con* and opportunity cost neglect. A wine trader has a frequent *buy* category that attends to the capital gain, so she exhibits less opportunity cost neglect (List [83]).

These predictions are based on the same mechanism: context and experiences prompt different categorizations of the same problem, focusing attention on different features and shaping choice. Problem recognition also explains non-fungibility: transferring money into a category prompts recognition of "buying in that category", promoting in-category spending. Mental categories also create commitment (Thaler [119]): setting up a "rainy day" account prompts recognition of "minimizing opportunity costs", raising similarity to *buy*, and discouraging spending. Account names focus the consumer on different goals, affecting how they recognize the problem and make their choices (Henderson and Peterson [66]).

4 Representations in Judgments

Expectations and probability judgments highlight the role of selective attention to different pieces of information. People with high inflation experiences attend more to inflation and its surges than people without such experiences (Gennaioli, Leva, Schoenle, and Shleifer [55], Link, Peichl, Pfauti, Roth and Wolfhart [82]). Professionals and households use narratives or models that attend to different data, and change over the cycle (Andre, Haaland, Roth, and Wohlfart [3], Bastianello, Décaire, and Guenzel [11]). Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20] show that in famous biases in i.i.d. draws and inference problems (Benjamin [15]) are also linked to changing attention to specific features.

These findings reject models of optimal beliefs or of stable heuristics/distortions of true probabilities (Grether [61]). We next show that this heterogeneity and instability emerges naturally from endogenous recognition of different statistical categories, unifying judgment and decision-making, and yielding new predictions.

4.1 Generalizing the Model

To allow for general DGPs, the event vector e is multi-dimensional. In balls and urns problems, one feature is urn selection $e_U \in \{urn\ A, urn\ B\}$, the other is the color of the ball drawn from the urn $e_B \in \{red, blue\}$. The event of selecting an urn and then drawing a ball from it is $e = (e_U, e_B)$, with probability $\mathbb{P}(e) = \mathbb{P}(e_B | e_U) \cdot \mathbb{P}(e_U)$. We generalize to multi-step causal models by an order $<$ over event features in Φ_E. The true probability of atom $y = (u, e)$ with $e = (e_s)_{s \in \Phi_E}$ is:

$$\mathbb{P}(e) = \prod_{s \in \Phi_E} \mathbb{P}\left(e_s | \cap_{j<s} e_j\right). \tag{15}$$

Attention shapes the DM's sensitivity to the probability of different events, as in Equation (2). Limited attention to event s, $\alpha_{P,s} < 1$, distorts its conditional probability into $\mathbb{P}\left(e_s | \cap_{j<s} e_j\right)^{\alpha_{P,s}}$. In the limit of zero attention, $\alpha_{P,s} = 0$, event s is "edited out", which prevents the conditioning of downstream events $s' > s$ on it. Formally, let edited vector $e(\alpha_P)$ report the correct event, $e_s(\alpha_P) = e_s$, if $\alpha_{P,s} > 0$ and set $e_s(\alpha_P) = \Omega$ otherwise, so that any realization of a neglected event is allowed.[14] The perceived probability Equation (15) is then:

[14] Observe that the discrepancy of how $\alpha_{P,s} = 0$ is treated here and in the single event feature environment of Section 2.1 is only apparent. Indeed, under both formulations $\Pr(e; \alpha_P)$ is a constant that does not affect the relative evaluation of alternatives.

$$\Pr(e;\alpha_P) \propto \prod_{s\in\Phi_E} \mathbb{P}\left(e_s | \cap_{j<s} e_j(\alpha_P)\right)^{\alpha_{P,s}}, \tag{16}$$

normalized so that probabilities across all e add to one. Equation (16) nests Grether's [61] formula, but here with endogenous weights. Valuation works as in Section 2: atom y is replaced by its edited version $y(\alpha_P) = (u(\alpha_P), e(\alpha_P))$ where $u(\alpha_P)$ satisfies (1). As in Equation (3), the value of $y(\alpha_P)$ multiplies its perceived probability and hedonics. The value $v(o;\alpha_P)$ of the edited lottery adds the values of its atoms. Full attention, $\alpha_P = \mathbf{1}$, recovers rationality.[15]

4.2 Biases in i.i.d. Draws and Inference

To jointly capture judgments about inference and i.i.d. sequences, consider a problem with $N+2$ event features. First, a coin (or urn) is selected of type $e_U \in \Theta$. This is followed by a sequence of N i.i.d. draws $(e_j)_{j\in\{1,...,N\}}$ in $\{h,t\}$ with outcomes *h*eads or *t*ails, whose probability is pinned down by e_U. The last feature is the sequence's share of heads $e_S \in \{x/N\}_{x\in\{0,1,...,N\}}$. Using Equation (16), the estimated probability of sequence e is given by:

$$\Pr(e;\alpha_P) \propto \mathbb{P}\left(e_S | e_N(\alpha_P), ..., e_U(\alpha_P)\right)^{\alpha_S} \prod_{i=1}^{N} \mathbb{P}\left(e_i | e_{i-1}(\alpha_P), ..., e_U(\alpha_P)\right)^{\alpha_i} \mathbb{P}(e_U)^{\alpha_U}, \tag{17}$$

which factorizes the distorted probabilities of the selected coin e_U, of its realized flips $(e_j)_{j\in\{1,...,N\}}$, and of the share of heads e_S. Full attention to these features gives a correct estimate.

We consider two categories based on the features we just described, capturing different statistical problems. Two context features play a central role: the coin selection and sequence length, denoted U and N respectively.

Frequency estimation. This category, $freq$, refers to experiences of estimating the probability of a single draw from a known process, e.g., the probability that a fair coin lands on h or t. Among the $N+2$ features in e, attention α_{freq} focuses on the i.i.d. draw corresponding to the hypothesis, for instance the first flip, $\alpha_1 = 1$ and $\alpha_i = 0$ for all other $i \in \Phi_E$. This category has two diagnostic context features:

[15] Here the decision rule computes the relative probability of hypotheses as the relative value of two lotteries (as in Savage [109], Section 9, we can allow for incentive compatible elicitation of multiple events). We can also apply our model to similarity judgments between objects A and B, as in Tversky's famous Austria example (Tversky [122], see Appendix B.4).

i) the coin is fair, $\kappa_{freq,U} = 0.5$ (denoting the probability of h) and ii) there is a single draw, $\kappa_{freq,N} = 1$.

Agnostic inference. Category inf refers to experiences based on i.i.d. noisy signals without having a DGP prior, for instance judging the quality of a restaurant based on a few reviews. Attention α_{inf} focuses on the share of positive signals (which in this case is a sufficient statistic for the DGP), setting $\alpha_S = 1$ and $\alpha_i = 0$ for all other $i \in \Phi_E$. The category also has two diagnostic context features: i) the coin type is not known, $\kappa_{inf,U} = [0, 1]$, and ii) there are at least two draws (selection of the DGP and one or more signals), $\kappa_{inf,N} \geq 2$.

When solving a statistical problem P, these categories compete for recognition, based on experiences and context. To illustrate, we show how the model generates the Gambler's Fallacy: people underestimate the probability of obtaining, with a fair coin, an unbalanced sequence such as *hhhhhh* relative to a balanced one of the same length, such as *htthht*. The model clarifies the conditions under which this bias arises, and makes new predictions. (We study biases in inference problems in Appendix B).

Current Problem P. The atoms of the i.i.d. problem we consider entail a *fair* coin, $e_U = 0.5$, but specify different flips e_j and share of heads e_S. The hypotheses are two specific atoms (here, the sequences), one balanced and the other not. The context vector κ_P reports that the coin is fair, $\kappa_{P,U} = 0.5$, and that the number of draws is $\kappa_{P,N} = N_P + 1$.[16]

The problem does not perfectly fit either category, $freq$ or inf, that the DM is familiar with. The coin is fair, $\kappa_{P,U} = 0.5$, which matches the frequency category but not inference, but there are many draws, which matches inference but not frequency.

Based on her categories, the DM can recognize P as $freq$ by focusing on the fair coin feature. This recognition draws her attention to the 50:50 nature of flips. Alternatively, the DM can recognize the problem as inf by focusing on the length of the sequence. This recognition draws her attention to the share of heads. Denote by d_N the similarity distance in the length of sequence feature N. With negligible shocks ϵ_c, $\lambda \to \infty$, we obtain the following result.

Proposition 5 *Suppose that $d_N (N_P + 1, \kappa_{inf,N}) = 0$ for all $N_P \geq 2$. There is a threshold $N (F_{freq}/F_{inf})$ with $N (\cdot)$ increasing in its argument, such that the DM recognizes P as $freq$ if $N_P < N (F_{freq}/F_{inf})$ and as inf otherwise. Estimated odds*

[16] As in Section 2, the context vector can include other features such as "the outcomes of the coin coincide with the hypotheses", or features of sequences such as "h and t are alternating". We abstract from these features because they do not play a role in this application.

satisfy:

$$\frac{v\left(htht....htht;\alpha_P^*\right)}{v\left(hhhh...hhhh;\alpha_P^*\right)}=\begin{cases}1 & \text{if } N_P<N\left(\frac{F_{freq}}{F_{inf}}\right)\\ \frac{\mathbb{P}(e_S=0.5|\Omega)}{\mathbb{P}(e_S=1|\Omega)}>1 & \text{if } N_P\geq N\left(\frac{F_{freq}}{F_{inf}}\right)\end{cases}. \tag{18}$$

The DM recognizes P as frequency estimation if N_P is low enough. She recognizes P as inference otherwise. Sequence length N_P shapes recognition because it is diagnostic of the two categories. Experiences also matter: higher relative familiarity with $freq$, higher F_{freq}/F_{inf}, favors this category over inference.

Critically, if N_P is low, the DM recognizes $freq$, she focuses on the fairness of the coin, and hence she correctly estimates the odds of sequences (despite her mis-estimation of their absolute probabilities).

In contrast, if N_P is high, then categorization switches to inf and the DM focuses on the share of heads, which is indeed relevant in inference. This causes her to represent P as assessing the "share of heads equivalence class" of *htht....htht* and *hhhh...hhhh*. This class is larger for balanced hypotheses, causing the GF. The inference representation causes the GF not because the DM believes that the coin type is not known. In fact, she uses the fair coin to compute the probability of e in Equation (16). It causes the GF because "inference" prompts the DM to focus on a feature of hypotheses, their share of heads, that causes question substitution. As a consequence, the DM makes an incorrect computation.

Our mechanism for the GF makes new predictions based on sequence length N_P and on other cognitive proxies. The first is experiences F_c: a DM who recently solved many inference problems should be more likely to commit the GF than one who has not, due to lower F_{freq}/F_{inf}. The second is salience: making the 50:50 nature of flips more prominent in the choice context should hinder the recognition of the *inf* category and the GF, as shown by Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20].[17] In Appendix B.2, we show that competition between the $freq$ and inf categories also accounts for multi-modality and instability in inference (Benjamin [15], Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20]).

We also shed light on the use of misspecified models neglecting the selection or correlation across sources of information (Enke and Zimmermann [45], Enke [43]). This evidence detects heterogeneity (some people account for correlation, others

[17] Rabin and Vayanos [99] show that a DM who incorrectly believes in the GF may exhibit a nonmonotonic pattern of under/overreaction depending on whether the DGP is independent or autocorrelated. Their model cannot explain why the same person exhibits the GF on one occasion but not on a normatively equivalent occasion based on changing context.

neglect it) and instability (neglect is stronger if sequences are longer, including based on redundant signals). Such biases emerge from recognition of inference, which focuses the DM on the share of heads, causing the neglect of the serial correlation. This account is consistent with the bias becoming: i) stronger with longer sequences and ii) weaker when selection in bottom-up salient. Individual level heterogeneity may in turn be due to experience with different problems.

This logic also sheds light on the use of narratives. People scarred by high inflation (Malmendier, Nagel, and Yan [91]) may view economic forecasting as "forecasting the costs of inflation", neglecting contemporaneous low unemployment. The same people can however focus on unemployment when it is salient in the media. When thinking about a shock to interest rates, households tend to focus on supply-side mechanisms, which are more familiar to them, but converge to a demand-side view when cued to think about business owners (Andre, Pizzinelli, Roth, and Wohlfart [4]). Investors may perceive an electric car company as belonging to the mature automotive sector, but become more optimistic if the firm reframes its business as "software", due to shifting attention across the firm's features (Sarkar [108]).

5 Bottom-up Attention and Framing

The allocation of attention is not determined exclusively top-down: it is also exogenously shaped by bottom-up salience. People attend more to known taxes when those appear on the price tag (Chetty, Looney, and Kroft [37]) and prefer goods that are physically present (Bushong, Camerer, and Rangel [29]). They also attend more to contrasting payoffs (Bordalo, Gennaioli, and Shleifer [21]-[22], Koszegi and Szeidl [77]).[18] The interaction of bottom-up salience and categorization yields a theory of framing: drawing attention to an already known feature of the problem can change how people solve it (Bordalo et al [20], Enke and Zimmermann [45]). These framing effects are readily testable because bottom-up salience can be experimentally controlled.

5.1 Bottom-up Salience

If a feature is shrouded, it has minimal bottom-up salience. If it is not shrouded, its salience is shaped by: i) sensory prominence and ii) contrast.

[18] See Lanzani [80] for the revealed preference axioms underpinning these effects.

Sensory prominence is captured in the description of P which, like the problem itself, is also an attention and context vector $(\alpha_\delta, \kappa_\delta)$. Subvector κ_δ reports the described context, which captures all that the DM perceives, so $\kappa_P = \kappa_\delta$. It is in part set by nature (e.g., "the sky" is exogenously part of κ_δ in an outdoor task), partly designed (e.g., a monetary lottery could be labeled as a "stock market bet"). Subvector $\alpha_\delta \in [0,1]^\Phi$ reports the prominence of different features in the description of P. If $\alpha_{\delta,i} = 0$ the feature i is shrouded as in Gabaix and Laibson [53]; if $\alpha_{\delta,i} = 1$ it is "front and center." Visual prominence can be measured as in Li and Camerer [81]; related methods can be used for text.

Contrast captures attention drawn to highly variable hedonics and events, as in Bordalo, Gennaioli, and Shleifer [22, 20]. Goal optimal attention (Sims [115], Woodford [127], Gabaix [52]) also depends on payoffs and probabilities, but contrast may *excessively* focus the DM on striking features, causing the neglect of others. Contrast of $i \in \Phi$ is a real valued function σ_i of that feature's realizations across atoms, $\sigma_i = \hat{\sigma}_i\left[(y_i)_{y\in Y}\right] \geq 1$, with $\hat{\sigma}_i\left[(y_i)_{y\in Y}\right] = 1$ if $y_i = y'_i$ for all $y, y' \in Y = \cup_{o\in O}$.[19]

Both the problem and the categories are sources of contrast. Contrast from the problem description, denoted $\sigma_{\delta,i}$, compares the values of feature i as described in P, e.g., the prices of different jams in the current supermarket. Contrast from category c, denoted $\sigma_{c,i}$, compares the values of i in experiences in c, such as variation in past jam prices. Often these measures overlap, but a feature can have a high category contrast – and be attended to – even if it is currently shrouded: before flying, some people think about crashes because they represent a contrasting payoff from memory. Contrast is computed using context vectors κ_δ and κ_c, which report the values of hedonic and event features.[20] Context features themselves have minimal contrast, because they do not vary across options.

Recognition. In Section 2, the DM sets attention to maximize the similarity

[19]A special contrast function is:

$$\hat{\sigma}_i\left[(y_i)_{y\in Y}\right] = 1 + \frac{\sum_{(y,y')\in Y^2: y\neq y'} f_i\left(y_i, y'_i\right)/|Y||Y-1|}{\sum_{y\in Y} f_i\left(y_i, \widetilde{y}_i\right)/|Y| + \varepsilon}, \tag{19}$$

where $\varepsilon > 0$ and $\widetilde{y}$ is a reference feature vector. For real-valued features, with $f_i\left(y_i, y'_i\right) = |y_i - y'_i|$ and $\widetilde{y} = 0$, it nests the exemplar function in Bordalo, Gennaioli, and Shleifer [22].

[20]If the feature is shrouded, $\alpha_{\delta,i} = 0$, its values are not encoded in κ_δ. In this case, description contrast is minimal, $\sigma_{\delta,i} = 1$. Similarly, category contrast is minimal, $\sigma_{c,i} = 1$, if the feature was neglected in c, $\alpha_{c,i} = 0$.

of the representation (α_P, κ_P) to the category (α_c, κ_c). Here, the DM allocates attention across features to trade off the problem's similarity to the category versus the description $(\alpha_\delta, \kappa_P)$. Equation (6) then becomes:

$$(\alpha_P^*, c^*) = \text{argmax}_{\alpha_P \in [0,1]^\Phi, c \in C} F_c \cdot S\left[(\alpha_P, \kappa_P), (\alpha_c, \kappa_c) | \sigma_c\right] + S\left[(\alpha_P, \kappa_P), (\alpha_\delta, \kappa_P) | \sigma_\delta\right] + \epsilon_c, \tag{20}$$

where the category-similarity function attaches higher weights to hedonics and events with more category contrast, and similarly for description-similarity (Nosofsky [97], Reed [101]):

$$S\left[(\alpha_P, \kappa_P), (\alpha_x, \kappa_x) | \sigma_x\right] \tag{21}$$
$$1 - \frac{\sum_{i \in \Phi} \sigma_{x,i} d\left(|\alpha_{P,i} - \alpha_{x,i}|\right) + \sum_{i \in \Phi_K} \alpha_{P,i} \alpha_{x,i} d_i\left(\kappa_{P,i}, \kappa_{x,i}\right)}{|\Phi| + |\Phi_K|} \qquad \forall x \in \{\delta, c\}.$$

We show the role of bottom-up salience using our consumption example (see also Proposition B.1 in Appendix B.1). We then shed light on well known framing effects in lottery choice.

5.2 Revisiting the Context Dependent Consumer

The riskless consumer problem studied in Section 3, based on neglect of the bad state b, raises two questions. First, what determines neglect vs. attention to the low-probability risk b? Unlikely events are often ignored, as in our example, but other times they are overweighted, as in Prospect Theory, Kahneman and Tversky [74]. Second, context is not the only driver of instability. Advertising and negotiation change choices by emphasizing some relevant features at the expense of others. How do such interventions work?

These questions are linked to bottom-up salience. Making a feature more bottom-up salient frames the problem in two ways: i) it affects attention to that feature within each category, and ii) it shapes categorization. We first explore how a frame shapes sensitivity to attributes *within* a category.

Let $\alpha_{P,i}(c|\delta)$ be the attention to feature i in a given category c when prominence is α_δ. This attention is found by maximizing similarity in Equation (20) at fixed c.

Proposition 6 *Attention $\alpha_{P,i}(c|\delta)$ to feature $i \in \Phi_H \cup \Phi_E$ increases in its prominence $\alpha_{\delta,i}$ and in category c's attention $\alpha_{c,i}$.*[21] *The description is more influential, i.e., $\inf_{\alpha_i \in \alpha_{P,i}(c|\delta)} |\alpha_i - \alpha_{\delta,i}|$ is lower, when $\sigma_{c,i} F_c / \sigma_{\delta,i}$ is lower. In particular, if*

[21] The increase in $\alpha_{P,i}(c|\delta)$ is with respect to the strong set order.

$d(x) = x$ *for all* $x \in \mathbb{R}_+$, *we have:*

$$\alpha_{P,i}(c|\delta) = \begin{cases} \alpha_{\delta,i} & if \quad \sigma_{c,i}F_c/\sigma_{\delta,i} < 1 \\ \alpha_{c,i} & if \quad \sigma_{c,i}F_c/\sigma_{\delta,i} > 1 \end{cases} \qquad \forall i \in \Phi_H \cup \Phi_E. \tag{22}$$

Increasing a feature's bottom up salience weakly increases the attention it receives independently of recognition. By Equation (22), this effect is strongest if $\sigma_{c,i}F_c/\sigma_{\delta,i} < 1$: feature i is attended to when it is prominent ($\alpha_{\delta,i} = 1$) and neglected when it is shrouded ($\alpha_{\delta,i} = 0$), even if the category prescribes otherwise. This occurs if the DM is either unfamiliar with c (F_c is low), or if the category contrast of feature i is small compared to its bottom up contrast ($\sigma_{c,i}/\sigma_{\delta,i}$ is low). Indeed, making quality prominent, $\alpha_{\delta,Qg} = 1$, promotes attention to it, especially if qualities on offer are contrasting, high $\sigma_{\delta,Qg}$, as in Bordalo, Gennaioli, and Shleifer [21, 22].

Conversely, bottom up salience does not matter if the category is a strong driver of attention to a feature, $\sigma_{c,i}F_c/\sigma_{\delta,i} > 1$. This can occur because the category is highly familiar, high F_c, or because the feature had high contrast in the past, high $\sigma_{c,i}/\sigma_{\delta,i}$. A person with many poverty experiences will attend to price even if it is shrouded, due to high F_{buy} and to past opportunity costs, high $\sigma_{c,M}$.

Bottom up salience sheds light on our first question, namely why are rare extreme payoffs sometimes neglected. This occurs when such payoffs are shrouded and have not been experienced in the past, in which case bottom-up salience and categorization jointly promote a riskless *con* or *buy* representation.

Changes in bottom up salience then shed light on our second question: why do representations sometimes change with irrelevant advertising and salience? This occurs when bottom up attention is *drawn* to previously neglected payoffs, particularly so if they are extreme, high $\sigma_{\delta,Qb}$. Crucially, this effect can cause a shift in categorization, giving a theory of framing effects. We illustrate this logic in our running example, and study the general case in the Appendix.[22]

Suppose that a batch of jam has been recalled due to contamination concerns. The media extensively describe the health risks, boosting the prominence $\alpha_{\delta,Qb}$ of the bad outcome q_b, which is severe compared to normal quality, $q_g+q_b < 0$. The problem

[22]Proposition B.1 in the Appendix shows this result holds in general. These forces help explain the experience-description gap in lottery choice (Hertwig and Erev [67]). They also produce the tendency to project today's tastes into the future (Loewenstein, O'Donoghue, and Rabin [85]). Bordalo, Gennaioli, and Shleifer [24] present a memory model in which the current experience helps simulate a similar future. Our framework suggests a related channel: the current experience is prominent so it receives a high weight in valuation.

can be represented as *consuming* a jam, or as *buying* it, or as "estimating the probability of a bad batch", using the $freq$ category. To see which representations prevails, assume that the context κ_{freq} associated to $freq$ in the past is distant from current context κ_P, so the consumer still chooses between *con* and *buy*, as in Section 3. Proposition 6 then implies the following result.

Proposition 7 *Suppose that context is sufficiently close to buying (κ_P low) that the consumer originally adopts a riskless buy category. As the prominence $\alpha_{\delta,Qb}$ of Qb increases from* 0 *to* 1*:*

i) Similarity of P to buy falls while that to con increases.

ii) Suppose that $d(x) = x$ for all $x \in [0,1]$, $\alpha_{\delta,W} = 0$, $\sigma_{con,W}F_{con}/\sigma_{\delta,W} < 1$, and that $\sigma_{con,i}F_{con}/\sigma_{\delta,i} > 1$ for all $i \in \Phi_H$. If the consumer switches to con, her willingness to pay falls to $(q_g + q_b)/2\alpha\eta$.

As in Proposition 6, making contamination concerns salient draws attention to the bad payoff in any category. Critically, and this is where reframing occurs, this description is more consistent with *con*, where qualities (good or bad) are fully attended to, than with *buy*, where quality shocks are neglected. The prominence of risk promotes re-categorization into *con*, leading to instability: not only does the consumer now focus on the downside q_b, but she *also* reduces her focus on price and neglects the fact that the probability $\mathbb{P}(b)$ is low. As a consequence, she exhibits: i) pessimistic beliefs about quality, and ii) a lower price sensitivity. Together, these imply that firms cannot prop up demand by cutting prices, even if by a lot.

A first key implication of Proposition 7 is to link information provision and framing. In a rational model, information about q_b causes a small drop in valuation, due to its small objective probability $\mathbb{P}(b)$. In Prospect Theory, a low $\mathbb{P}(b)$ is exaggerated, but price sensitivity does not vary with information.[23] In our model, in contrast, information causes instability of attention across a *range* of features, including the downside risk, its probability, but also monetary incentives. Information provision affects decisions more through changing context than its news content.

The well known "certainty effect" (Kahneman and Tversky's [74]) is an instance of this mechanism: adding a rare but high-contrast downside risk discontinuously reduces the valuation of a sure thing (Barseghyan, Molinari, O'Donoghue, and Teitelbaum [10], Haigh and List [63]) because it directs attention toward extreme payoffs,

[23]Proposition 7 implies that recategorization requires a sufficiently large increase in the salience of risks. People often shrug when told about risks, so long as they are not made prominent enough.

not toward small probabilities.[24,25]

A second implication of Proposition 7 is that the problem's description α_δ matters if it raises the prominence of a feature that is "important", in the sense that it is either attended to in a frequent category or that it varies sharply across options. Framing effects are not "random psychology" or residual noise. They rely on problem recognition and, as such, are the strongest when they change the salience of important features.

5.3 Instability in Lottery Choice

A large body of work on lottery choice in the lab demonstrates systematic instability of risk attitudes relative to the expected utility benchmark, typically explained using Prospect Theory's reference points or probability weighting functions (Kahneman and Tversky [74]). Our model accounts for these anomalies, revealing deep connections between different choice domains – probability judgments and choice among goods – and yielding new predictions. We illustrate these ideas with several examples, leaving a full treatment to future work.

Lottery choices in lab experiments entail prominent payoffs and probabilities. The DM is torn between a "payoff evaluation" or a "frequency estimation" problem. The first representation is tied to the *con* category, and attention to payoff tradeoffs. The second is tied to the $freq$ category, and attention to probability tradeoffs. Experiences in *con* are arguably more frequent, but they occur in settings that are less similar to the abstract context of the lab. Instability in categorization, due to variations in bottom-up salience, entails instability of risk attitudes.

Payoff Prominence. Risk attitudes change with lottery descriptions even if payoffs and probabilities are held constant. Suppose a DM chooses between a lottery paying $m_g > 0$ with probability $\mathbb{P}(g)$, and $m_b = 0$ otherwise, and a sure thing paying its expected value $m = m_g \cdot \mathbb{P}(g)$, which delivers the same payoff m in states g and

[24] When choosing between \$100 and \$99 categorization into *con* correctly leads to the choice of \$100. Adding a small risk ε of a \$0 payoff, to the \$100 payoff introduces a downside feature with contrasting values \$0 and \$99 across options. As it draws attention, it reinforces categorization into *con*, prompting neglect of probabilities and risk aversion. Bordalo, Gennaioli, and Shleifer [22] study the role of payoff contrast but their continuous payoff weights foreclose discontinuities.

[25] Another important example of discontinuity arises in situations involving social norms. Gneezy and Rustichini [57] show that a small payment reduces effort in the collection of donations, presumably because the payment is now categorized as a low-salary job.

b. Consider two descriptions of the risky lottery.

$$\text{Shrouded Downside} \quad : \quad \text{win } m_g \text{ with probability } \mathbb{P}(g). \tag{23}$$

$$\text{Full Prominence} \quad : \quad \text{win } m_g \text{ with probability } \mathbb{P}(g) \text{ and 0 otherwise.} \tag{24}$$

In Equation (23), the downside payoff and its probability are implied but not prominently described, $\alpha_{\delta,M_b} < 1$ and $\alpha_{\delta,W} < 1$. In Equation (24), the prominence of the downside payoff increases to $\alpha_{\delta,M_b} = 1$. This has two effects. First, it increases attention to the downside in any category (with a logic similar to the one of Proposition 6). Second, it promotes adoption of *con* (Proposition 7). As a result, in Equation (4) the DM attaches a higher weight α_{P,M_b} to the downside payoff for both the lottery and the sure thing, which favors the safe choice. A mere increase in descriptive prominence of the downside risk increases risk aversion even though the lottery has not changed.[26]

This effect of prominence is absent from existing theories of choice under risk but can be important in the real world, where prominence varies across payoffs and over time. To take a few examples, advertising high financial returns promotes investment and the neglect of risk, especially when the crash event is shrouded (Mullainathan, Schwartzstein, and Shleifer [95], Célérier and Vallée [32]).[27] Advertising top prizes drives the demand for lottery tickets (Kearney [87]). Demand for insurance coverage rises after uninformative but superficially similar shocks (Kunreuther [78], Dessaint and Matray [40]). Different parties to a transaction or negotiation emphasize gains or losses differently (Kahneman and Lovallo [72]).

Contrast of payoffs and probabilities. We next illustrate how changing payoffs or probabilities can change their contrast, shift attention across features, and cause re-categorization. Consider the Allais common ratio effect (Allais [1], Kahneman and Tversky [74]). If people are indifferent between \$100 with probability 0.2 and \$20 with probability 0.8, then they likely prefer \$100 with probability 0.02 to \$20 with

[26]Formally, valuation of lottery l and of sure thing s in category c are, respectively:

$$\begin{aligned}\tilde{v}(l,c) &= \mathbb{P}(g)^{\alpha_{P,W}(c)} \cdot \alpha_{P,M_g}(c) \cdot m_g, \\ \tilde{v}(s,c) &= \left[\mathbb{P}(g)^{\alpha_{P,W}(c)} \cdot \alpha_{P,M_g}(c) + (1-\mathbb{P}(g))^{\alpha_{P,W}(c)} \cdot \alpha_{P,M_b}(c)\right] \cdot m.\end{aligned}$$

Higher attention to the downside payoff $\alpha_{P,M_b}(c)$ boosts the valuation of the sure thing, not of the lottery, promoting the safe choice.

[27]A related treatment varies the description of the sure thing by explicitly mentioning that it never yields a downside (unlike the risky alternative), or by keeping it implicit. We thank Alex Imas for making this point as a discussant.

probability 0.08. This pattern violates expected utility, where preferences should be invariant to uniform scaling of probabilities.

Our model applies as follows. There is a good lottery state g and a bad state b. The two lotteries deliver different payoffs in g, creating payoff contrast $\sigma_{\delta,Mg}$. They also differ in the probability of state g, creating event contrast $\sigma_{\delta,e}$. Crucially, event contrast changes with the scaling of probabilities: $\sigma_{\delta,e}$ is high when comparing probabilities 0.2 and 0.8, but declines dramatically when comparing 0.02 to 0.08. Proposition B.1 in Appendix B.1 shows that this decline in $\sigma_{\delta,e}$ reduces attention to probabilities, and promotes a shift towards the "payoff estimation" representation *con*, which also increases attention to payoffs. This change in weights benefits the riskier lottery, generating the common ratio Allais paradox.[28]

When differences in probabilities are large the DM attends to them and neglects payoffs, encouraging risk aversion. When differences in probabilities are "peanuts" the DM neglects them and attends to payoffs, encouraging risk taking. This shift in representations offers a foundation for Rubinstein's [104] intuition that choice depends on perceived similarity between probabilities.

Instability in risk attitudes driven by probability contrast illuminates risky choices in the field. For instance, in both baseball and basketball players routinely decide whether to favor riskier, high payoff shots (home-runs and 3-pointers respectively). But while home-runs were overvalued in baseball prior to Moneyball (Hakes and Sauer [65]), basketball players have been reluctant to take 3-point shots despite their higher expected value (Schwartzstein and Malhotra [112]).

The Allais common ratio logic can shed light on this behavior. In basketball, the safer 2-pointer has a hit rate of 50% compared to 35% for the 3-pointer. In baseball, the safer "walk" has a hit rate of 9% compared to 3.5% for the home run. While in basketball two pointer is $50/35 = 1.43$ times more likely to succeed than a three pointer, in baseball a walk is $9/3.5 = 2.57$ times more likely to succeed than a home run. At given payoffs this logic favors more risk taking in basketball than in baseball, contrary to historical practice.

[28]Denote the safer lottery by $(\mathbb{P}(w), w)$ and the mean preserving spread by $(\mathbb{P}(x), x)$. Then, the valuation gap $\widetilde{v}(c)$ between the lotteries in category c is given by (see Appendix B.3):

$$\widetilde{v}(c) = \mathbb{P}(x)^{\alpha_{P,e}(c)} \cdot \alpha_{P,u}(c) \cdot (x - w) + \left[\mathbb{P}(x)^{\alpha_{P,e}(c)} - \mathbb{P}(w)^{\alpha_{P,e}(c)}\right] \cdot w \cdot \alpha_{P,u}(c).$$

Since $\mathbb{P}(x) \cdot x = \mathbb{P}(w) \cdot w$, the right hand side increases in attention to payoffs $\alpha_{P,u}(c)$ and decreases in attention to probabilities $\alpha_{P,e}(c)$. In particular, categorization in *con* boosts risk taking while categorization in $freq$ boosts risk aversion.

Our model explains the observed historical practice because the *difference* in hit rates is much larger in basketball than in baseball (15 versus 5.5 percentage points). Larger probability differences draw attention in basketball, favoring the safer 2-pointer, while small probability differences are neglected in baseball, favoring the home run.[29]

Lastly, in Appendix B.3, we show that categorization into *con* due by high contrast payoff differences also explains the fourfold pattern in lotteries, as well as in riskless mirrors (Oprea [98], de Clippel, Oprea, and Rozen [39]).[30] A common cognitive mechanism of attention driven by categorization and bottom-up salience unifies biases in riskless choice, probability judgments, and risky choice.

6 Conclusion

Canonical models of decision making, both neoclassical and behavioral, hinge on the maximization of a stable objective function that combines hedonics and probabilities. We present a new approach in which the DM first draws on selective attention and memory to mentally represent the problem at hand, and then combines the hedonics and probabilities that are relevant in that representation. As a consequence, the objective function that the DM maximizes exhibits spurious heterogeneity and instability based on past experiences, context, and framing. In contrast to both neoclassical and standard behavioral models, this approach unifies observed heterogeneity and instability in judgments and decision making.

Our analysis raises two important methodological issues. First, to properly understand behavior, cognitive forces such as attention and memory must be formalized and measured. Choice analysis, even if enriched with psychological carriers of utility, is insufficient. Our model makes precise predictions based on cognitive inputs such as context, experiences, and similarity (see Bordalo, Conlon, Gennaioli, Kwon,

[29]Payoffs move in the same direction of making the risky strategy more appealing in baseball than basketball. One home-run is worth between 4 and 16 walks, which is much larger than the difference between 2 and 3 points in basketball.

[30]For similar reasons, our model helps explain weak within-person correlation of choices both within a domain, e.g., low correlation between insurance demand and lottery choice (Barseghyan, Prince, and Teitelbaum [9]) and across domains, e.g., low correlation between the endowment effect and aversion to mixed lotteries (Chapman, Dean, Ortoleva, Snowberg, and Camerer [35]). Although in this paper we limit ourselves to the case of known probabilities, the same mechanism could connect the striking relation between behavior facing ambiguity and decisions with known probabilities documented in Halevy [64].

and Shleifer [20], Bordalo, Burro, Coffman, Gennaioli and Shleifer [17], Malmendier and Nagel [90]). It also opens the door for a structured use of AI methods, which can recover representations from self-reported reasons for choice (Haaland, Roth, Stantcheva, and Wohlfart [62], Link, Peschl, Pfauti, Roth, and Wohlfart [82]), as well as unveil subtle contextual features as in Ludwig and Mullainathan [86]. Future work can use such data in a theoretically disciplined way to study the cognitive structure of decisions.

The second methodological point concerns experimental design. Current practice favors the use of abstract protocols to identify "universal" choice biases and minimize experimenter demand. Abstraction can help capture general cognitive forces, but is problematic for studying real-world choices, such as demand for insurance, for two reasons. First, removing naturalistic context can also remove consumers' spontaneous representations, reducing external validity for predicting field behavior. Second, elaborate and abstract protocols geared at measuring "true" tastes may anchor assessments to laboratory context features, providing a distorted metric of normative decision quality. "Naturalistic immersion" is a benefit of field experiments, but our theory suggests that future work should also develop methods to properly engineer controlled variation of naturalistic context in the lab.

Our analysis opens several new avenues, and we conclude with two of them. The first concerns the nature and origin of categories. We showed that two riskless choice and two statistical categories suffice to unify a great deal of evidence. They could also help explain evidence we did not discuss. For instance, by focusing people on the cost of giving away resources, the *buy* category may help account for phenomena related to loss aversion, such as reluctance to take the risk of losing money or the unwillingness to pay to acquire a good.[31] Of course, there are many more categories. In risky choice, a natural category can entail "regret" or "elation" in which attention is focused on these features of experience. In intertemporal choice, a natural category is one of "investment" decisions, in which attention is focused on long-term payoffs. Future work should study where such categories come from. One possibility is that they are formed through bottom-up salient features of experiences, such as a feeling a striking payoff or being exposed to social reminders of cultural norms, which then become diagnostic markers for different categories of problems. Because such bottom-up forces are potentially measurable, endogenous categories would enhance the explanatory power of our approach.

[31] In this way, we could endogenize the utility function in Rabin and Weizsäcker [100], which combines integrated and nonintegrated lotteries, as a function of experience and description.

A second major avenue is to study the real-world implications of our mechanism. When thinking about redistribution, some voters may think about fairness, others about zero-sum transfers from taxpayers (Chinoy, Nunn, Sequeira, and Stantcheva [38]) based on different experiences, but changes in context such as the specific name or domain of the tax may change problem similarity, representations, and voter preferences. Some people think guns at home bring safety, others that they bring the risk of an accident (Alsan, Schwartzstein, and Stantcheva [2]). Similar considerations apply to fairness judgments (Kahneman, Knetsch, and Thaler [71]), which may be viewed from the perspective of equity or of endowment; to strategic behavior (List [84]), which may be viewed in terms of maximizing self-interest or punishing transgressive behavior; and many other settings.

Ongoing work explores the role of reasoning in real world settings that also require estimation of payoffs and probabilities. Bordalo et al. (2026) show that priming individuals to recall personal adversity shifts their thinking about inflation away from economic news and toward concerns about affordability and personal distress, leading them to expect higher inflation. Such real world applications are a key testing ground not only of the model's explanatory power but also of the design of choice architectures to improve mental representations and choices.

Supplementary Material

An Online Appendix for this article can be found at The Quarterly Journal of Economics online.

References

[1] Allais, Maurice, 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine. Econometrica 21(4), pp. 503-546.

[2] Alsan, Marcella, Schwartzstein, Joshua, Stantcheva, Stefanie, 2026. The universal pursuit of safety and the demand for (lethal, non-lethal or no) guns. Mimeo.

[3] Andre, Peter, Roth, Christopher, Haaland, Ingar and Wohlfart, Johannes, 2021. Inflation narratives.

[4] Andre, Peter, Pizzinelli, Carlo, Roth, Christopher and Wohlfart, Johannes, 2022. Subjective models of the macroeconomy: Evidence from experts and representative samples. The Review of Economic Studies, 89(6), pp.2958-2991.

[5] Awh, Edward, Belopolsky, Artem, and Theeuwes, Jan, 2012, Top-down versus bottom-up attentional control: a failed theoretical dichotomy, Trends in Cognitive Science, 16, 437-443.

[6] Ba, Cuimin, Bohren, J. Aislinn and Imas, Alex, 2024. Over-and underreaction to information. SSRN 4274617.

[7] Baicker, Katherine, Mullainathan, Sendhil, and Schwartzstein, Joshua, 2015. Behavioral hazard in health insurance. The Quarterly Journal of Economics, 130(4), pp.1623-1667.

[8] Banerjee, Abhijit and Duflo, Esther, 2007. The economic lives of the poor. Journal of Economic Perspectives, 21, pp.141-167.

[9] Barseghyan, Levon, Prince, Jeffrey, and Teitelbaum, Joshua, 2011. Are risk preferences stable across contexts? Evidence from insurance data. American Economic Review, 101(2), 591-631.

[10] Barseghyan, Levon, Molinari, Francesca, O'Donoghue, Ted and Teitelbaum, Joshua, 2013. The nature of risk preferences: Evidence from insurance choices. American Economic Review, 103(6), pp.2499-2529.

[11] Bastianello, Francesco, Décaire, Paul, and Guenzel, Marius, 2024. Mental models and financial forecasts. SSRN working paper 5004839.

[12] Bassok, Miriam, 2003. Analogical transfer in problem solving. In J. E. Davidson and R. J. Sternberg (Eds.), The psychology of problem solving (pp. 343-369). Cambridge University Press.

[13] Becker, Gary, 1962. Irrational behavior and economic theory. Journal of Political Economy, 70(1), pp. 1-13.

[14] Becker, Gary and Murphy, Kevin, 1993. A simple theory of advertising as a good or bad. The Quarterly Journal of Economics, 108(4), pp. 941-964.

[15] Benjamin, Daniel, 2019. Errors in probabilistic reasoning and judgment biases. Handbook of Behavioral Economics: Applications and Foundations 1, 2, pp.69-186.

[16] Bohren, J. Aislinn, Imas, Alex, Ungeheuer, Michael and Weber, Martin, 2024. A Cognitive Foundation for Perceiving Uncertainty. SSRN 32149.

[17] Bordalo, Pedro, Burro, Giovanni, Coffman, Katherine, Gennaioli, Nicola and Shleifer, Andrei, 2025. Imagining the future: memory, simulation and beliefs about COVID. The Review of Economic Studies, 92 (3), pp. 1532-1563.

[18] Bordalo, Pedro, Burro, Giovanni, Gennaioli, Nicola, Nacamulli, Guido, and Shleifer, Andrei, 2025. Ads as Cues. National Bureau of Economic Research, w34387.

[19] Bordalo, Pedro, Conlon, John, Gennaioli, Nicola, Kwon, Spencer, and Shleifer, Andrei, 2023. Memory and probability. The Quarterly Journal of Economics 138 (1): 265-311.

[20] Bordalo, Pedro, Conlon, John, Gennaioli, Nicola, Kwon, Spencer, and Shleifer, Andrei, 2026. How people use statistics. The Review of Economic Studies, 93 (1): 250-285.

[21] Bordalo, Pedro, Gennaioli, Nicola and Shleifer, Andrei, 2012. Salience theory of choice under risk. The Quarterly Journal of Economics, 127(3), pp.1243-1285.

[22] Bordalo, Pedro, Gennaioli, Nicola and Shleifer, Andrei, 2013. Salience and consumer choice. Journal of Political Economy, 121(5), pp.803-843.

[23] Bordalo, Pedro, Gennaioli, Nicola and Shleifer, Andrei, 2016. Competition for attention. The Review of Economic Studies 83 (2): 481-513.

[24] Bordalo, Pedro, Gennaioli, Nicola and Shleifer, Andrei, 2020. Memory, attention, and choice. The Quarterly Journal of Economics, 135(3), pp.1399-1442.

[25] Bordalo, Pedro, Gennaioli, Nicola and Shleifer, Andrei, 2022. Salience. Annual Review of Economics, 14, pp.521-544.

[26] Bordalo, Pedro, Gennaioli, Nicola, La Porta, Rafael, and Shleifer, Andrei, 2024. Belief overreaction and stock market puzzles. Journal of Political Economy 132 (5), pp. 1450-1484.

[27] Bordalo, Pedro, Gennaioli, Nicola, La Porta, Rafael and Shleifer, Andrei, 2025. Finance without exotic risk. Journal of Financial Economics, 173, pp.104-145.

[28] Bordalo, Pedro, Gennaioli, Nicola, Lopez de Silanes, Florencio, Schroder, Steffen, Shleifer, Andrei, and van Rooij, Maarten, 2026. The psychology of macroeconomic expectations. Working paper.

[29] Bushong, Benjamin, King, Logan, Camerer, Colin and Rangel, Antonio, 2010. Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay. American Economic Review, 100 (4), pp.1556-1571.

[30] Camerer, Colin, and Johnson, Eric, 1991. The process performance paradox in expert judgment. In A. Ericsson & J. Smith (Eds.), Toward a General Theory of Expertise.

[31] Camerer, Colin, Johnson, Eric, Rymon, Talia, and Sen, Sankar, 1993. Cognition and framing in sequential bargaining for gains and losses. In Frontiers of Game Theory, 27â€“47. Cambridge, MA: MIT Press.

[32] Célérier, Claire and Vallée, Boris, 2017. Catering to investors through security design: Headline rate and complexity. The Quarterly Journal of Economics, 132(3), pp.1469-1508.

[33] Cenzon, Josefina, 2025. Credit market experiences and macroeconomic expectations: Evidence and theory. Working paper.

[34] Chandra, Amitabh, Flack, Evan, Obermeyer, Ziad, 2024. The health costs of cost sharing. The Quarterly Journal of Economics, 139 (4), 2037 - 2082.

[35] Chapman, Jonathan, Dean, Mark, Ortoleva, Pietro, Snowberg, Erik and Camerer, Colin, 2024. Willingness to accept, willingness to pay, and loss aversion. National Bureau of Economic Research, w30836.

[36] Chase, William, and Simon, Herbert, 1973. Perception in chess. Cognitive Psychology, 4(1), pp. 55-81.

[37] Chetty, Raj, Looney, Adam and Kroft, Kory, 2009. Salience and taxation: Theory and evidence. American Economic Review, 99(4), pp.1145-1177.

[38] Chinoy, Sahil, Nunn, Nathan, Sequeira, Sandra and Stantcheva, Stefanie, 2023. Zero-sum thinking and the roots of US political divides, American Economic Review 116 (3), pp. 1052â€“96.

[39] de Clippel, Geoffroy, Oprea, Ryan, and Rozen, Kareen, 2025. As If, working paper.

[40] Dessaint, Olivier, and Matray, Adrien, 2017. Do Managers Overreact to Salient Risks? Evidence from Hurricane Strikes. Journal of Financial Economics 126(1), pp.97 - 121.

[41] Edwards, Ward, 1968. Conservatism in Human Information Processing. in Kahneman, Daniel, Slovic, Paul and Tversky, Amos (eds) Judgment under Uncertainty: Heuristics and Biases (Cambridge: Cambridge University Press), pp. 359-369.

[42] Ellis, Andrew and Masatlioglu, Yusufcan, 2022. Choice with endogenous categorization. The Review of Economic Studies, 89(1), pp.240-278.

[43] Enke, Benjamin, 2020. What you see is all there is. The Quarterly Journal of Economics, 135(3), pp.1363-1398.

[44] Enke, Benjamin and Graeber, Thomas, 2023. Cognitive uncertainty. The Quarterly Journal of Economics, 138(4), pp.2021-2067.

[45] Enke, Benjamin and Zimmermann, Florian, 2019. Correlation neglect in belief formation. The Review of Economic Studies, 86(1), pp.313-332.

[46] Fan, Tingzhe, Liang, Yueran, and Peng, Cameron, 2026. The Inference-Forecast Gap in Belief Updating. Econometrica, forthcoming.

[47] Festinger, Leon, 1957. A Theory of Cognitive Dissonance. Evanston, IL: Row and Peterson.

[48] Frederick, Shane, Novemsky, Nathan, Wang, Jing, Dhar, Ravi and Nowlis, Stephen, 2009. Opportunity cost neglect. Journal of Consumer Research, 36(4), pp.553-561.

[49] Fryer, Roland and Jackson, Matthew, 2008. A categorical model of cognition and biased decision making. The BE Journal of Theoretical Economics, 8(1).

[50] Fudenberg, Drew, Lanzani, Giacomo and Strack, Philipp, 2024. Selective memory equilibrium. Journal of Political Economy, 132, pp. 3978-4020.

[51] Fudenberg, Drew, Lanzani, Giacomo and Strack, Philipp, 2025. Learning, Memory, and Stochastic Choice. Available at SSRN 4923469.

[52] Gabaix, Xavier, 2019. Behavioral inattention. In Handbook of behavioral economics: Applications and foundations 1 (Vol. 2, pp. 261-343). North-Holland.

[53] Gabaix, Xavier and Laibson, David, 2006. Shrouded attributes, consumer myopia, and information suppression in competitive markets. The Quarterly Journal of Economics, 121(2), pp.505-540.

[54] Gagnon-Bartsch, Tristan, Rabin, Matthew and Schwartzstein, Joshua, 2023. Channeled attention and stable errors, Mimeo.

[55] Gennaioli, Nicola, Leva, Marta, Schoenle, Raphael and Shleifer, Andrei, 2024. How inflation expectations de-anchor: The role of selective memory cues. National Bureau of Economic Research, w32633.

[56] Gilboa, Itzhak and Schmeidler, David, 1995. Case-based decision theory. The Quarterly Journal of Economics, 110(3), pp.605-639.

[57] Gneezy, Uri and Rustichini, Aldo, 2000. Pay enough or don't pay at all. The Quarterly Journal of Economics, 115(3), pp.791-810.

[58] Goldstone, Robert, 1984. Influences of categorization on perceptual discrimination. Journal of Experimental Psychology: General 123 (2): 178.

[59] Gourville, John T. and Soman, Dilip, 1998. Payment depreciation: The behavioral effects of temporally separating payments from consumption. Journal of Consumer Research 25(2), pp. 160-174.

[60] Green, Etan, Rao, Justin and Rothschild, David, 2019. A sharp test of the portability of expertise. Management Science, 65(6), pp.2820-2831.

[61] Grether, David M., 1980. Bayes rule as a descriptive model: The representativeness heuristic. The Quarterly Journal of Economics, 95(3), pp.537-557.

[62] Haaland, Ingar, Roth, Christopher, Stantcheva, Stefanie and Wohlfart, Johannes, 2024. Measuring what is top of mind. National Bureau of Economic Research, w32421.

[63] Haigh, Michael and List, John, 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. Journal of Finance, 60(1), pp.523-534.

[64] Halevy, Yoram, 2007. Ellsberg revisited: An experimental study. Econometrica, 75(2), pp.503-536.

[65] Hakes, Jahn, and Sauer, Raymond, 2006. An economic evaluation of the Moneyball hypothesis. Journal of Economic Perspectives, 20(3), pp. 173 - 186.

[66] Henderson, Pamela and Peterson, Robert, 1992. Mental accounting and categorization. Organizational Behavior and Human Decision Processes, 51(1), pp.92-117.

[67] Hertwig, Ralph and Erev, Ido, 2009. The descriptionâ€“experience gap in risky choice. Trends in cognitive sciences, 13(12), pp.517-523.

[68] Hoch, Stephen, Kim, Byung-Do, Montgomery, Alan and Rossi, Peter, 1995. Determinants of store-level price elasticity. Journal of Marketing Research, 32(1), pp.17-29.

[69] Itti, Laurent and Baldi, Pierre, 2009. Bayesian surprise attracts human attention. Vision Research, 49(10), pp.1295-1306.

[70] Jehiel, Philippe, 2005. Analogy-based expectation equilibrium. Journal of Economic Theory, 123(2), pp.81-104.

[71] Kahneman, Daniel, Knetsch, Jack and Thaler, Richard, 1986. Fairness as a constraint on profit seeking: Entitlements in the market. American Economic Review, 76(4), pp.728-741.

[72] Kahneman, Daniel and Lovallo, Dan, 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. Management Science, 39(1), pp.17-31.

[73] Kahneman, Daniel and Tversky, Amos, 1972. Subjective probability: A judgment of representativeness. Cognitive Psychology, 3, pp. 430-454.

[74] Kahneman, Daniel and Tversky, Amos, 1979. Prospect Theory: an analysis of decision under risk. Econometrica, 47(2), pp.263-292.

[75] Kahneman, Daniel and Tversky, Amos, 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90(4), pp. 293 - 315.

[76] Kőszegi, Botond and Matějka, Filip, 2020. Choice simplification: A theory of mental budgeting and naive diversification. The Quarterly Journal of Economics, 135(2), pp.1153-1207.

[77] Kőszegi, Botond and Szeidl, Adam, 2013. A model of focusing in economic choice. The Quarterly Journal of Economics, 128(1), pp.53-104.

[78] Kunreuther, Howard, 1978. Disaster insurance protection: Public policy lessons. New York: John Wiley & Sons.

[79] Laibson, David, 2001. A cue-theory of consumption. The Quarterly Journal of Economics, 116(1), pp.81-119.

[80] Lanzani, Giacomo, 2022. Correlation made simple: Applications to salience and regret theory. The Quarterly Journal of Economics, 137(2), pp.959-987.

[81] Li, Xiaomin and Camerer, Colin F., 2022. Predictable effects of visual salience in experimental decisions and games. The Quarterly Journal of Economics, 137(3), pp.1849-1900.

[82] Link, Sebastian, Peichl, Andreas, Pfauti, Oliver, Roth, Christopher, and Wohlfart, Johannes, 2025. Attention to the Macroeconomy. SSRN n. 4697814.

[83] List, John, 2004. Neoclassical theory versus prospect theory: Evidence from the marketplace. Journal of Political Economy, 72(2), pp.615-625.

[84] List, John, 2007. On the interpretation of giving in dictator games. Journal of Political Economy, 115(3), pp.482-493.

[85] Loewenstein, George, O'Donoghue, Ted, and Rabin, Matthew, 2003. Projection bias in predicting future utility, The Quarterly Journal of Economics, 118(4), pp. 1209 - 1248

[86] Ludwig, Jens and Mullainathan, Sendhil, 2024. Machine learning as a tool for hypothesis generation. The Quarterly Journal of Economics, 139(2), pp.751-827.

[87] Kearney, Melissa, 2005. State lotteries and consumer behavior. Journal of Public Economics, 89(11-12), pp.2269-2299.

[88] Mack, Michael and Palmeri, Thomas J., 2024. Discrimination, Recognition, and Classification, The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications.

[89] Malmendier, Ulrike, 2021. Experience effects in finance: Foundations, applications, and future directions. Review of Finance, 25, pp. 1339-1363.

[90] Malmendier, Ulrike and Nagel, Stefan, 2016. Learning from inflation experiences. The Quarterly Journal of Economics, 131(1), pp.53-87.

[91] Malmendier, Ulrike, Nagel, Stefan, and Yan, Zhen, 2021. The Making of Hawks and Doves: Inflation Experiences on the FOMC. Journal of Monetary Economics, 117, pp. 19-42

[92] Manski, Charles, 1995. Identification problems in the social sciences. Cambridge, MA: Harvard University Press.

[93] Mohlin, Erik, 2014. Optimal categorization. Journal of Economic Theory, 152, pp.356-381.

[94] Mullainathan, Sendhil, 2002. Thinking through categories, Working Paper, Harvard University.

[95] Mullainathan, Sendhil, Schwartzstein, Joshua and Shleifer, Andrei, 2008. Coarse thinking and persuasion. The Quarterly Journal of Economics, 123(2), pp.577-619.

[96] Nickerson, Raymond S., 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), pp.175-220.

[97] Nosofsky, Robert, 1986. Attention, similarity, and the identification–categorization relationship. Journal of Experimental Psychology: General, 115(1), p.39.

[98] Oprea, Ryan, 2024. Decisions under risk are decisions under complexity. American Economic Review, 114(12), pp.3789-3811.

[99] Rabin, Matthew and Vayanos, Dimitri, 2010. The gambler's and hot-hand fallacies: Theory and applications. The Review of Economic Studies, 77(2), pp.730-778.

[100] Rabin, Matthew and Weizsäcker, Georg, 2009. Narrow bracketing and dominated choices. American Economic Review, 99(4), pp.1508-1543.

[101] Reed, Stephen K., 1972. Pattern recognition and categorization. Cognitive Psychology, 3(3), pp.382-407.

[102] Rick, Scott, Cryder, Cynthia, and Loewenstein, George, 2008. Tightwads and spendthrifts. Journal of Consumer Research, 34(6), pp.767-782.

[103] Rosch, Eleanor and Lloyd, Barbara, 1978. Principles of Categorization, Routledge.

[104] Rubinstein, Ariel, 1988. Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). Journal of Economic Theory, 46(1), pp.145-153.

[105] Rumelhart, David, and McClelland, James, 1981. An interactive activation model of context effects in letter perception. Psychological Review, 88(5), pp. 375-407.

[106] Salant, Yuval and Rubinstein, Ariel, 2008. (A, f): choice with frames. The Review of Economic Studies, 75(4), pp.1287-1296.

[107] Samuelson, Larry, 2001. Analogies, adaptation, and anomalies. Journal of Economic Theory, 97(2), pp.320-366.

[108] Sarkar, Shraman, 2025. Economic representations, Mimeo.

[109] Savage, Leonard, 1971. Elicitation of personal probabilities and expectations. Journal of the American Statistical Association, 66(336), pp.783-801.

[110] Schwartzstein, Joshua, 2014. Selective attention and learning. Journal of the European Economic Association, 12(6), pp.1423-1452.

[111] Schwartzstein, Joshua and Sunderam, Adi, 2021. Using models to persuade. American Economic Review, 111(1), pp.276-323.

[112] Schwartzstein, Joshua, and Malhotra, Deepak, 2021. Rocket Science. Harvard Business School Case 921-043, October 2021. Boston: Harvard Business School Publishing.

[113] Shafir, Eldar and Thaler, Richard, 2006. Invest now, drink later, spend never: On the mental accounting of delayed consumption. Journal of Economic Psychology, 27(5), pp.694-712.

[114] Shah, Anuj, Zhao, Jiaying, Mullainathan, Sendhil and Shafir, Eldar, 2018. Money in the mental lives of the poor. Social Cognition, 36(1), pp.4-19.

[115] Simon, Herbert, 1955. A behavioral model of rational choice. The Quarterly Journal of Economics, pp.99-118.

[116] Sims, Christopher, 2003. Implications of rational inattention. Journal of Monetary Economics, 50(3), pp.665-690.

[117] Stigler, George and Becker, Gary, 1977. De gustibus non est disputandum. The American Economic Review, 67(2), pp.76-90.

[118] Taubinsky, Dmitry, Butera, Luigi, Saccarola, Matteo and Lian, Chen, 2025. Beliefs About the Economy are Excessively Sensitive to Household-Level Shocks: Evidence from Linked Survey and Administrative Data. National Bureau of Economic Research, w32664.

[119] Thaler, Richard, 1980. Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization, 1(1), pp.39-60.

[120] Thaler, Richard, 1999. Mental accounting matters. Journal of Behavioral Decision Making, 12(3), pp.183-206.

[121] Tversky, Amos, 1972. Elimination by aspects: A theory of choice. Psychological Review, 79(4), p.281 - 299.

[122] Tversky, Amos, 1977. Features of similarity. Psychological Review, 84(4), p.327 - 352.

[123] Tversky, Amos and Kahneman, Daniel, 1974. Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), pp.1124-1131.

[124] Tversky, Amos and Kahneman, Daniel, 1981. The framing of decisions and the psychology of choice. Science, 211(4481), pp.453-458.

[125] Wachter, Jessica, and Kahana, Michael, 2024. A retrieved-context theory of financial decisions. The Quarterly Journal of Economics, 139(2), pp.1095-1147.

[126] Wakefield, Kirk, and Inman, Jeffrey, 2003. Situational price sensitivity: the role of consumption occasion, social context and income. Journal of Retailing, 79(4), pp.199-212.

[127] Woodford, Michael, 2012. Prospect theory as efficient perceptual distortion. American Economic Review, 102(3), pp.41-46.

Online Appendix

for

A Cognitive Theory of Reasoning and Choice

by

Pedro Bordalo Nicola Gennaioli
Giacomo Lanzani Andrei Shleifer

A Categorization

A.1 Notation Index

Symbol	Meaning
O	Set of lotteries that can be chosen
$\left(\Omega, 2^{\Omega}, \mathbb{P}\right)$	Probability space (known to the DM)
Φ_H	Set of hedonic features of the problem
Φ_K	Set of context features of the problem
κ	Context vector
$P = \{O, \kappa_P\}$	Problem
$u = (u_i)_{i\in\Phi_H} \in \mathbb{R}^{\Phi_H}$	Vector of hedonic features
Φ_E	Set of event features of the problem
$\Phi = \Phi_H \cup \Phi_E \cup \Phi_K$	Set of features of the problem
$e = (e_i)_{i\in\Phi_E}$	Vector of event features
$y = (u, e)$	Atom of a lottery
$\alpha_P \in [0,1]^{\Phi}$	Attention vector
$\overline{u}_i = \frac{\sum_{o\in O}\sum_{(u,e)\in o} u_i}{\sum_{o\in O} \lvert o\rvert}$	Average value of u_i across all atoms
$u_i\left(\alpha_P\right) = \alpha_{P,i}\cdot u_i + \left(1-\alpha_{P,i}\right)\cdot \overline{u}_i$	Attention-based value of hedonic i
$e_i\left(\alpha_P\right)$	Attention-based value of event i
$y\left(\alpha_P\right) = \left(u\left(\alpha_P\right), e\left(\alpha_P\right)\right)$	Attention-based perception of atom y
$v\left(y;\alpha_P\right)$	Attention-based evaluation of atom y
$v\left(o;\alpha_P\right) = \sum_{y\in o} v\left(y;\alpha_P\right)$	Attention-based evaluation of lottery o
$>$	Linear causal order over event features
$o\left(\alpha_P\right)$	Attention-based perception of lottery o
Z	Set of possible situations
d_i	Distance metric in context feature $i \in \Phi_K$
C	Set of categories
F_c	Temporally discounted frequency of category c

Symbol	Meaning
S	Similarity function
$d : \mathbb{R}_+ \to \mathbb{R}$	Distance between attention levels
$\lambda \in \mathbb{R}_{++}$	Scale parameter in the similarity perception noise

A.2 Proofs of Section 2

Proof of Proposition 1. First of all, observe that if $d_i(\kappa_{P,i},\kappa_{c,i})\,\mathbb{I}_{\{1\}}(\alpha_{c,i})\,\mathbb{I}_{\{\Phi_K\}}(i)=0$, the minimal distance of 0 in feature i can (only) be achieved by letting $\alpha_{P,i}(c)=\alpha_{c,i}$, so clearly the weight fully follows the category ones. Moreover, if $i\in\Phi_K$, $\alpha_{c,i}=1$, and $d_i(\kappa_{P,i},\kappa_{c,i})>0$, then any $\alpha_{P,i}(c)\in[0,1]$ weakly shrinks attention towards 0.

We are left to establish that $d_i(\kappa_{P,i},\kappa_{c,i})\geq d_i\left(\kappa'_{P,i},\kappa'_{c,i}\right)$ implies

$$\begin{aligned}&\left\{\bar{\alpha}_{P,i}\in[0,1]:\bar{\alpha}_P\in\operatorname{argmax}_{\alpha_P\in[0,1]^\Phi} S\left[(\alpha_P,\kappa_P),(\alpha_c,\kappa_c)\right]\right\}\\ \leq\ _{SSO}&\left\{\bar{\alpha}_{P,i}\in[0,1]:\bar{\alpha}_P\in\operatorname{argmax}_{\alpha_P\in[0,1]^\Phi} S\left[(\alpha_P,\kappa'_P),(\alpha_c,\kappa'_c)\right]\right\}.\end{aligned}$$

It is trivial that if $\alpha_{c,i}=0$, both sets coincide with $\{0\}$, so consider the case where $\alpha_{c,i}=1$. Observe that $S\left[(\alpha_P,\kappa_P),(\alpha_c,\kappa_c)\right]$ depends on $\alpha_{P,i}$ and $d_i(\kappa_{P,i},\kappa_{c,i})$ only additively through the term

$$d(\alpha_{c,i}-\alpha_{P,i})+\alpha_{P,i}d_i(\kappa_{P,i},\kappa_{c,i})\,\alpha_{c,i}=d(1-\alpha_{P,i})+\alpha_{P,i}d_i(\kappa_{P,i},\kappa_{c,i})\,.$$

If we see this as a family $\{d(1-(\cdot))+(\cdot)\,d_i(\kappa_{P,i},\kappa_{c,i})\}_{d_i(\kappa_{P,i},\kappa_{c,i})}$ of functions of $\alpha_{P,i}$ parametrized by $d_i(\kappa_{P,i},\kappa_{c,i})$, it clearly satisfies single crossing differences, and therefore the statement follows by Milgrom and Shannon [4]. ■

Lemma 1 *Let $\alpha_P(c)$ be an arbitrary element of* $\operatorname{argmax}_{\alpha_P\in[0,1]^\Phi} S\left[(\alpha_P,\kappa_P),(\alpha_c,\kappa_c)\right]$. *The maximum total similarity satisfies:*

$$\partial S(P,c)/\partial F_c \;=\; 1-d(P,c)\geq 0, \tag{25}$$

$$\partial S(P,c)/\partial d_i(\kappa_{P,i},\kappa_{c,i}) \;\propto\; -F_c\alpha_{P,i}(c)\leq 0. \tag{26}$$

Proof. Observe that

$$d(P,c)=1-\max_{\alpha_P\in[0,1]^\Phi} S\left[(\alpha_P,\kappa_P),(\alpha_c,\kappa_c)\right].$$

Then it follows that

$$S(P,c)=F_c(1-d(P,c)) \tag{27}$$

and Equation (25) follows immediately.

Finally,

$$\frac{\partial S(P,c)}{\partial d_i\left(\kappa_{P,i},\kappa_{c,i}\right)} = F_c \frac{\partial \max_{\alpha_P\in[0,1]^\Phi} S\left[\left(\alpha_P,\kappa_P\right),\left(\alpha_c,\kappa_c\right)\right]}{\partial d_i\left(\kappa_{P,i},\kappa_{c,i}\right)} = F_c\frac{-\alpha_{P,i}(c)\alpha_{c,i}}{|\Phi|+|\Phi_K|} = F_c\frac{-\alpha_{P,i}(c)}{|\Phi|+|\Phi_K|} \leq 0$$

where the second step follows from the envelope theorem (see Theorem 2 of Milgrom and Segal [3] for a result that can be applied here), and the third from the fact that if $\alpha_{c,i}=0$, then by Proposition 1 we also have $\alpha_{P,i}(c)=0$. ■

Proof of Proposition 2. Equation (9) follows by, e.g., Lemma 1 in McFadden [2] and the usual observations that: (1) if ϵ_c is Gumbel with precision parameter λ then $\lambda\epsilon_c$ is Gumbel with precision parameter 1 and (2) $F_c\left(1-d\left(P,c\right)\right)+\epsilon_c \geq F_{c'}\left(1-d\left(P,c'\right)\right)+\epsilon_{c'}$ if and only if $\lambda F_c\left(1-d\left(P,c\right)\right)+\lambda\epsilon_c \geq \lambda F_{c'}\left(1-d\left(P,c'\right)\right)+\lambda\epsilon_{c'}$. Then, using Equation (9) we calculate

$$\begin{aligned}\frac{\partial \Pr\left(c|P\right)}{\partial S(P,c)} &= \frac{\left(\sum_{c'\in C}\exp\left[\lambda S\left(P,c'\right)\right]\right)\lambda\exp\left[\lambda S\left(P,c\right)\right]-\exp\left[\lambda S\left(P,c\right)\right]\lambda\exp\left[\lambda S\left(P,c\right)\right]}{\left(\sum_{c'\in C}\exp\left[\lambda S\left(P,c'\right)\right]\right)^2}, \\ &= \lambda\left[\frac{\exp\left[\lambda S\left(P,c\right)\right]}{\sum_{c'\in C}\exp\left[\lambda S\left(P,c'\right)\right]}-\left(\frac{\exp\left[\lambda S\left(P,c\right)\right]}{\sum_{c'\in C}\exp\left[\lambda S\left(P,c'\right)\right]}\right)^2\right] \\ &= \lambda\left[\Pr\left(c|P\right)-\left(\Pr\left(c|P\right)\right)^2\right]=\lambda\Pr\left(c|P\right)\left[1-\Pr\left(c|P\right)\right].\end{aligned}$$

Combining this with Lemma 1 (in particular, Equations (25) and (26)) and using the chain rule immediately yields the desired result. ■

Let $o_{Pc}\in O$ be the option with the highest evaluation conditional on adopting category c, which is assumed for semplicity to be unique and with $o_{Pc}\neq o_{Pc'}$ whenever $c\neq c'$.

Lemma 2 *For all $c\in C$:*
i) Higher F_c increases $\Pr\left(o\left(c^*\right)=o_{Pc}\right)$ and weakly decreases $\Pr\left(o\left(c^*\right)=o_{Pc'}\right)$ for all $c'\neq c$. Take two DMs j and j' with $\sum_{c'\in C}F_{c'}^j=\sum_{c'\in C}F_{c'}^{j'}$. They choose o_{Pc} with different probabilities for some $c\in C$ if and only if $F_c^j\neq F_c^{j'}$ for some $c\in C$.
ii) Higher $d(P,c)$ decreases $\Pr\left(o\left(c^*\right)=o_{Pc}\right)$ and weakly increases $\Pr\left(o\left(c^*\right)=o_{Pc'}\right)$ for all $c'\neq c$.
iii) Decreasing $d(P,c)$, boosts $\Pr\left(o\left(c^*\right)=o_{Pc}\right)$ more at higher F_c if and only if c is not dominant, i.e., $F_c<F_c^*$ where the threshold F_c^* increases in $d(P,c)$. In particular, this always holds if $\Pr(c|P)\leq 1/2$.

Proof. Observe preliminarily that $d(P,c) < 1$ for all $c \in C$, as by setting $\alpha_{P,i^*} = \alpha_{c,i^*}$ for some $i^* \in I$ the DM

$$\frac{\sum_{i\in\Phi\setminus\{i^*\}} d(|\alpha_{c,i} - \alpha_{P,i}|) + \sum_{i\in\Phi_K} \alpha_{P,i}\alpha_{c,i}d_i(\kappa_{P,i},\kappa_{c,i})}{|\Phi| + |\Phi_K|} \le \frac{|\Phi| + |\Phi_K| - 1}{|\Phi| + |\Phi_K|} < 1.$$

For (i), note that by Proposition 2,

$$\frac{\partial \Pr(o(c^*) = o_{Pc})}{\partial F_c} = \frac{\partial \Pr(c|P)}{\partial F_c} > 0,$$

so choice of o_{Pc} is more likely when F_c is higher.

Now, consider two different individuals. We first recall that by Proposition 2,

$$\Pr(c^* = c\,|P,j) = \frac{\exp\left(\lambda F_c^j(1 - d(P,c))\right)}{\sum_{c'\in C}\exp\left(\lambda F_{c'}^j(1 - d(P,c'))\right)}.$$

Observe that $d(P,c)$ is constant across individuals, and only depends on the category c. It is immediately clear that if $F_c^j = F_c^{j'}$ for all $c \in C$, then $\Pr(c^* = c\,|P,j) = \Pr(c^* = c\,|P,j')$ for all $c \in C$. So we just need to prove the converse, which is that if $\Pr(c^* = c\,|P,j) = \Pr(c^* = c\,|P,j')$ for all $c \in C$, then $F_c^j = F_c^{j'}$ for all $c \in C$.

For sake of contradiction, assume that we can select some $\hat{c} \in C$ with $F_{\hat{c}}^j \ne F_{\hat{c}}^{j'}$. Without loss of generality, let $F_{\hat{c}}^j > F_{\hat{c}}^{j'}$. For every $c' \in C \setminus \{\hat{c}\}$ we have

$$\begin{aligned}\frac{\Pr(c^* = \hat{c}\,|P,j)}{\Pr(c^* = c'\,|P,j)} &= \frac{\exp\left(\lambda F_{\hat{c}}^j(1 - d(P,\hat{c}))\right)}{\exp\left(\lambda F_{c'}^j(1 - d(P,c'))\right)} \\ &= \exp\left(\lambda\left[F_{\hat{c}}^j(1 - d(P,\hat{c})) - F_{c'}^j(1 - d(P,c'))\right]\right).\end{aligned}$$

Therefore

$$\begin{aligned}&\frac{\Pr(c^* = \hat{c}\,|P,j)}{\Pr(c^* = c'\,|P,j)} = \frac{\Pr(c^* = \hat{c}\,|P,j')}{\Pr(c^* = c'\,|P,j')} \\ &\implies F_{\hat{c}}^j(1 - d(P,\hat{c})) - F_{c'}^j(1 - d(P,c')) \\ &= F_{\hat{c}}^{j'}(1 - d(P,\hat{c})) - F_{c'}^{j'}(1 - d(P,c')) \\ &\implies \left(F_{\hat{c}}^j - F_{\hat{c}}^{j'}\right)(1 - d(P,\hat{c})) = \left(F_{c'}^j - F_{c'}^{j'}\right)(1 - d(P,c'))\end{aligned}$$

so $F_{\hat{c}}^{j} > F_{\hat{c}}^{j'}$ means that

$$F_{c'}^{j} > F_{c'}^{j'} \qquad \forall c' \in C \setminus \{\hat{c}\}.$$

But this contradicts $\sum_{c' \in C} F_{c'}^{j} = \sum_{c' \in C} F_{c'}^{j'}$.

For (ii), recall from the proof of Proposition 2 that

$$\frac{\partial \Pr(c|P)}{\partial S(P,c)} = \lambda \Pr(c|P)\left[1 - \Pr(c|P)\right],$$

and also the fact that, by Equation (27), $\frac{\partial S(P,c)}{\partial d(P,c)} = -F_c$, we have

$$\frac{\partial \Pr(c|P)}{\partial d(P,c)} = -\lambda \cdot F_c \cdot \Pr(c|P) \cdot \left[1 - \Pr(c|P)\right].$$

Then applying Proposition 2 again, we get

$$\begin{aligned}
\frac{\partial^2 \Pr(c|P)}{\partial d(P,c)\,\partial F_c} &= -\lambda \Pr(c|P)\left[1 - \Pr(c|P)\right] - \lambda F_c \frac{\partial \Pr(c|P)}{\partial F_c}\left[1 - \Pr(c|P)\right] + \lambda F_c \Pr(c|P)\frac{\partial \Pr(c|P)}{\partial F_c} \\
&\propto -\Pr(c|P)\left[1 - \Pr(c|P)\right] - \lambda F_c (1 - d(P,c))\left[1 - \Pr(c|P)\right]\Pr(c|P)\left[1 - \Pr(c|P)\right] \\
&\quad + \lambda \cdot F_c \cdot (1 - d(P,c)) \cdot \Pr(c|P)\Pr(c|P) \cdot \left[1 - \Pr(c|P)\right] \\
&= \left(-1 + \lambda S(P,c)\left[2\Pr(c|P) - 1\right]\right)\Pr(c|P) \cdot \left[1 - \Pr(c|P)\right].
\end{aligned}$$

So

$$\frac{\partial^2 \Pr(c|P)}{\partial d(P,c)\,\partial F_c} \leq 0 \iff \lambda F_c (1 - d(P,c))\left[2\Pr(c|P) - 1\right] \leq 1.$$

The condition holds for F_c small enough (and whenever $\Pr(c|P) \leq 1/2$), but does not hold for F_c large enough, which in particular guarantees $\Pr(c|P) > 1/2$ (from Equation (9) the latter condition holds provided $\exp(\lambda F_c (1 - d(P,c))) > \sum_{c' \neq c} \exp(\lambda F_{c'} (1 - d(P,c')))$). Therefore, since $\lambda F_c (1 - d(P,c))\left[2\Pr(c|P) - 1\right]$ is increasing in F_c and decreasing in $d(P,c)$ in the range $\Pr(c|P) \geq 1/2$, there exists a threshold F_c^*, which increases in $d(P,c)$, such that the condition holds if and only if $F_c < F_c^*$, establishing (iii).

For every $\alpha \in [0,1]^{\Phi}$, let α_H denote the restriction of α to the hedonic features Φ_H.

■

Lemma 3 *Suppose that there are no nontrivial event features and that $\alpha_P(c)$ is a singleton for all $c \in C$. Let*

$$\overline{\alpha}_i = \sum_{c\in C} \alpha_{P,i}(c) \Pr(c|P) \qquad \forall i \in \Phi.$$

denote the average attention profile. For every $o \in O$ with hedonic vector u, the average valuation

$$\overline{v(o)} = \sum_{c\in C} \Pr(c|P)\, v(o; \alpha_P(c))$$

satisfies:

$$\frac{\partial \overline{v(o)}}{\partial F_c} = \lambda (1 - d(P,c)) \Pr(c|P) \langle \alpha_{P,H}(c) - \overline{\alpha_H}, u - \bar{u}\rangle, \quad \forall c \in C \tag{28}$$

$$\frac{\partial \overline{v(o)}}{\partial d_i(\kappa_{P,i}, \kappa_{c,i})} = \frac{-\lambda F_c \alpha_{P,i}(c) \Pr(c|P) \langle \alpha_{P,H}(c) - \overline{\alpha_H}, u - \bar{u}\rangle}{|\Phi| + |\Phi_K|} \quad \forall c \in C, \forall i \in \Phi_K \tag{29}$$

Proof. Let $\overline{\alpha}_{P-c} = \frac{\overline{\alpha}_P - \alpha_P(c)\Pr(c|P)}{[1-\Pr(c|P)]}$ be the average attention weights conditional on not being categorized in c. We have that

$$\begin{aligned}
\frac{\partial \overline{v(o)}}{\partial F_c} &= \langle \alpha_{P,H}(c)\partial \Pr(c|P)/\partial F_c - \overline{\alpha}_{P-c,H}(c)\partial \Pr(c|P)/\partial F_c, u - \bar{u}\rangle \\
&= \left\langle \alpha_{P,H}(c)\partial \Pr(c|P)/\partial F_c - \frac{\overline{\alpha}_H - \alpha_{P,H}(c)\Pr(c|P)}{[1-\Pr(c|P)]}\partial \Pr(c|P)/\partial F_c, u - \bar{u}\right\rangle \\
&= \frac{\partial \Pr(c|P)}{\partial F_c}\left\langle \frac{\alpha_{P,H}(c) - \overline{\alpha}_H}{[1-\Pr(c|P)]}, u - \bar{u}\right\rangle = \lambda(1 - d(P,c))\Pr(c|P)\langle \alpha_{P,H}(c) - \overline{\alpha}_H, u - \bar{u}\rangle
\end{aligned}$$

where the fourth equality follows by Proposition 2, proving Equation (28). The proof of Equation (29) is completely analogous. ■

A.3 Proofs for Section 3

Proof of Proposition 3. 1. That $(\alpha_{P,Qg}(buy), \alpha_{P,Mg}(buy)) = (\alpha, 1)$ follows from Proposition 1. Then, observe that the derivative of the similarity with respect to $\alpha_{P,Z}$ is $\frac{1}{8} - \frac{1}{4}\kappa_P$ which is strictly positive if $\kappa_P < \frac{1}{2}$ and strictly negative if $\kappa_P > \frac{1}{2}$. The formula for $S(P, buy)$ follows by plugging in Equation (10) these optimal values.

2. That $(\alpha_{P,Qg}(con), \alpha_{P,Mg}(con)) = (1, \alpha)$ follows from Proposition 1. Then, observe that the derivative of the similarity with respect to $\alpha_{P,Z}$ is $\frac{1}{8} - \frac{1}{4}(1-\kappa_P)$ which is strictly positive if $\kappa_P > \frac{1}{2}$ and strictly negative if $\kappa_P < \frac{1}{2}$. The formula for $S(P, con)$ follows by plugging in Equation (10) these optimal values. ■

Proof of Proposition 4. By Proposition 3, category *buy* is adopted if

$$\kappa_P < \frac{7F_{buy}}{2F_{con}} - 3 \text{ and } \kappa_P > \frac{1}{2}$$

or

$$\kappa_P < 4 - \frac{7}{2}\frac{F_{con}}{F_{buy}} \text{ and } \kappa_P < \frac{1}{2}.$$

Since

$$4 - \frac{7}{2}\frac{F_{con}}{F_{buy}} < \frac{1}{2} \Longleftrightarrow \frac{F_{con}}{F_{buy}} > 1 \Longleftrightarrow \frac{7F_{buy}}{2F_{con}} - 3 < \frac{1}{2},$$

we have that category *buy* is adopted if $\frac{F_{con}}{F_{buy}} > 1$ and $\kappa_P < 4 - \frac{7}{2}\frac{F_{con}}{F_{buy}}$ or $\frac{F_{con}}{F_{buy}} < 1$ and $\kappa_P < \frac{7F_{buy}}{2F_{con}} - 3$ that is

$$\kappa\left(\frac{F_{buy}}{F_{con}}\right) = \begin{cases} 4 - \frac{7}{2}\frac{F_{con}}{F_{buy}} & \text{if } \frac{F_{con}}{F_{buy}} > 1 \\ \frac{7F_{buy}}{2F_{con}} - 3 & \text{if } \frac{F_{con}}{F_{buy}} < 1 \end{cases}$$

which is increasing in $\frac{F_{buy}}{F_{con}}$. ■

Proof of Proposition 5. First observe that upon adopting the frequency category, by Proposition 1 the only event feature to which the decision maker assigns positive weight is the first flip, which results in an evaluation ratio equal to 1. Analogously, upon adopting the agnostic inference category, by Proposition 1 the only event feature to which the decision maker assigns positive weight is the share of heads, which results in an evaluation ratio equal to $\frac{\mathbb{P}(e_S=0.5|\Omega)}{\mathbb{P}(e_S=1|\Omega)} > 1$.

Next, fix a frequency ratio F_{freq}/F_{inf}. Suppose that if $N_P = N^*$ the DM categorizes the problem as agnostic inference. We show that she also categorizes the problem as inference when $N_{P'} = N^* + 2$ (observe that the constraint of a 50:50 share of heads implies that we only consider the case of an even $N_P, N_{P'}$). By Proposition 1, both with N_P and $N_{P'}$ the attention to the number of flips upon matching to the agnostic inference category is set to 1. Therefore, we have the desired

$$d\left(P', inf\right) = d\left(P, inf\right)\frac{N^*+6}{N^*+8} \leq d\left(P, freq\right)\frac{N^*+6}{N^*+8} \leq d\left(P', freq\right),$$

where the last inequality follows from the fact that $d\left(P, freq\right)$ is computed with respect to the optimal allocation of attention over the N^*+6 feature of problem P. Finally, that the threshold is increasing in $\frac{F_{freq}}{F_{inf}}$ immediately follows by Lemma 1. ■

References

[1] List, John, Muir, Ian, Pope, Devin and Sun, Gregory, 2023. Left-digit bias at lyft. The Review of Economic Studies, 90(6), pp.3186-3237.

[2] McFadden, Daniel, 1974. Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics.

[3] Milgrom, Paul and Segal, Ilya, 2002. Envelope theorems for arbitrary choice sets. Econometrica, 70(2), pp.583-601.

[4] Milgrom, Paul and Shannon, Chris, 1994. Monotone comparative statics. Econometrica: Journal of the Econometric Society, pp.157-180.

B Bottom-up Attention

B.1 Proofs for Section 5

Proof of Proposition 6. Let $i \in \Phi_H \cup \Phi_E$ and define $\sigma^* = \frac{\sigma_{\delta,i}}{\sigma_{c,i}} \cdot \frac{1}{F_c}$. Since the absolute value is convex, and d is convex and nondecreasing, $S\left(P, c, \delta|\sigma\right)$ is a concave function of $\alpha_{P,i}$. This, inter alia, implies that $\alpha_{P,i}(c|\delta)$ is always an interval. Moreover, the continuity of the absolute value and d imply that the interval is closed. Observe that $F_c S\left[\left(\alpha_P, \kappa_P\right), \left(\alpha_c, \kappa_c\right)|\sigma_c\right] + S\left[\left(\alpha_P, \kappa_P\right), \left(\alpha_\delta, \kappa_P\right)|\sigma_\delta\right]$ depends on $\alpha_{P,i}$, $\alpha_{c,i}$, and $\alpha_{\delta,i}$ only additively through the term

$$d\left(|\alpha_{P,i} - \alpha_{c,i}|\right) + \sigma^* d\left(|\alpha_{P,i} - \alpha_{\delta,i}|\right). \tag{30}$$

If we see this as a family $\{d(|(\cdot)-\alpha_{c,i}|)+\sigma^* d(|(\cdot)-\alpha_{\delta,i}|)\}_{\alpha_{c,i}\in[0,1]}$ of functions of $\alpha_{P,i}$ parametrized by $\alpha_{c,i}$, it satisfies single crossing differences by the monotonicity and convexity of d, so $\alpha_{P,i}(c|\delta)$ is increasing (with respect to the strong set order) in $\alpha_{c,i}$ by Milgrom and Shannon [4]. Similarly, when we see Equation (30) as a family

$$\{d(|(\cdot)-\alpha_{c,i}|)+\sigma^* d(|(\cdot)-\alpha_{\delta,i}|)\}_{\alpha_{\delta,i}\in[0,1]}$$

of functions of $\alpha_{P,i}$ parametrized by $\alpha_{\delta,i}$, it satisfies single crossing differences by the monotonicity and convexity of d, so $\alpha_{P,i}(c|\delta)$ is increasing (with respect to the strong set order) in $\alpha_{\delta,i}$ by Milgrom and Shannon [4].

For the second part of the proposition, suppose first that $\alpha_{\delta,i}\in\alpha_{P,i}(c|\delta,\sigma^*)$, where we abused the notation and made the dependence of $\alpha_{P,i}(c|\delta)$ on σ^* explicit. It is immediate to see from Equation (30) that if σ^* is increased to some $(\sigma^*)'$ with $\sigma<(\sigma^*)'$, $\alpha_{\delta,i}\in\alpha_{P,i}(c|\delta,(\sigma^*)')$, as desired. Next, consider first the case in which we start from a situation where

$$\underline{\alpha}_i :=\min_{\alpha\in\alpha_P(c|\delta,\sigma^*)}\alpha_i>\alpha_{\delta,i}.$$

It is immediate that in this case we must have $\alpha_{c,i}\geq\underline{\alpha}_i$. But then, for every $\alpha_i'\geq\underline{\alpha}_i$, the optimality condition for $\underline{\alpha}_i$ under σ^*,

$$\begin{aligned} & d(|\alpha_i'-\alpha_{c,i}|)+\sigma^* d(|\alpha_i'-\alpha_{\delta,i}|) \\ \geq\ & d(|\underline{\alpha}_i-\alpha_{c,i}|)+\sigma^* d(|\underline{\alpha}_i-\alpha_{\delta,i}|) \end{aligned}$$

implies, together with the monotonicity of d, that for $(\sigma^*)'\geq\sigma^*$

$$\begin{aligned} & d(|\alpha_i'-\alpha_{c,i}|)+(\sigma^*)' d(|\alpha_i'-\alpha_{\delta,i}|) \\ \geq\ & d(|\underline{\alpha}_i-\alpha_{c,i}|)+(\sigma^*)' d(|\underline{\alpha}_i-\alpha_{\delta,i}|) \end{aligned}$$

proving that

$$\alpha_{\delta,i}\leq\min_{\alpha\in\alpha_P(c|\delta,(\sigma^*)')}\alpha_i\leq\underline{\alpha}_i$$

as desired. The proof for the case where $\max_{\alpha\in\alpha_{P,i}(c|\delta,\sigma^*)}\alpha_i<\alpha_{\delta,i}$ is completely symmetric and thus omitted.

The last conclusion for the case where $d(x)=x$ for all $x\in\mathbb{R}_+$ immediately follows by inspection of Equation (30). ■

Critically, description shapes - via attention - similarity and categorization. To see this, denote by $S(P,c|\delta)$ be the maximum of Equation (20) for a given category c.

Proposition B.1 *Let d be differentiable, $i \in \Phi_H \cup \Phi_E$, and $\alpha_{\delta,i} > \alpha_{c,i}$. Increasing $\alpha_{\delta,i}$ or $\sigma_{\delta,i}$ increases similarity more for categories in which that feature is more relevant. Formally, $\frac{\partial S(P,c|\delta)}{\partial \alpha_{\delta,i}}$ and $\frac{\partial S(P,c|\delta)}{\partial \sigma_{\delta,i}}$ are increasing in $\alpha_{c,i}$.*

Proof. Let $i \in \Phi_H \cup \Phi_E$ and define $\sigma^* = \frac{\sigma_{\delta,i}}{\sigma_{c,i} F_c}$. Since the absolute value is convex, and d is convex and nondecreasing, $S(P, c, \delta|\sigma)$ is a concave function of $\alpha_{P,i}$. This, inter alia, implies that $\alpha_{P,i}(c|\delta, \sigma)$ is always an interval (we slightly abuse notation to make explicit the dependence of the solution on δ). Moreover, the continuity of the absolute value and d imply that the interval is closed. Thus let $\alpha_P^{\min}(c|\delta, \sigma) = \text{argmin}_{\alpha \in \alpha_P(c|\delta,\sigma)} \alpha$.[32] We can compute $\frac{\partial S(P,c,\delta|\sigma)}{\partial \alpha_{c,i}}$ for $i \in \Phi_H \cup \Phi_E$ using the envelope theorem (see, e.g., Milgrom and Segal [3], Theorems 1 and 3) to obtain

$$\frac{\partial S(P, c, \delta|\sigma)}{\partial \alpha_{c,i}} = -\sigma_{c,i} F_c \frac{\frac{\partial}{\partial \alpha_{c,i}} d(|\alpha_{P,i}^{\min}(c|\delta, \sigma) - \alpha_{c,i}|)}{|\Phi| + |\Phi_K|}.$$

Since, by Proposition 6, $\alpha_{P,i}^{\min}(c|\delta, \sigma)$ is increasing in $\alpha_{\delta,i}$ and d is convex, $\frac{\partial S(P,c,\delta|\sigma)}{\partial \alpha_{c,i}}$ is increasing in $\alpha_{\delta,i}$. Thus, straightforward computations give that $\frac{\partial S(P,c,\delta|\sigma)}{\partial \alpha_{\delta,i}}$ is increasing in $\alpha_{c,i}$.

Since, by Proposition 6, $\alpha_{P,i}^{\min}(c|\delta, \sigma)$ is increasing in $\sigma_{\delta,i}$ and d is convex, $\frac{\partial S(P,c,\delta|\sigma)}{\partial \alpha_{c,i}}$ is increasing in $\sigma_{\delta,i}$. Thus, straightforward computations give that $\frac{\partial S(P,c,\delta|\sigma)}{\partial \sigma_{\delta,i}}$ is increasing in $\alpha_{c,i}$. ■

Categories focused on prominent and contrasting features are more likely to be selected. This is a theory of framing effects: a bottom-up salient feature can produce a re-categorization of the problem, causing preference reversals.

Proof of Proposition 7. Observe that given the additively separable structure of the similarity function, changes in α_{δ,Q_b} only change similarity through changes in the term

$$F_c \sigma_{c,Qb} d(|\alpha_{P,Qb} - \alpha_{c,Qb}|) + \sigma_{\delta,Qb} d(|\alpha_{P,Qb} - \alpha_{\delta,Qb}|). \tag{31}$$

We first prove part (i). For $c \in \{buy, freq\}$, when $\alpha_{\delta,Qb} = 0$ the term in Equation (31) has $\alpha_{P,Qb}$ optimally set to 0 so that $d(|\alpha_{P,Qb} - \alpha_{c,Qb}|) = 0 = d(|\alpha_{P,Qb} - \alpha_{\delta,Qb}|)$, as $\alpha_{c,Qb} = \alpha_{\delta,Qb} = 0$. Therefore, this

[32] Observe this minimization is vectorial, but the additive separability of the problem guarantees the existence of an α which solves jointly the minimization across all dimensions.

term can only increase (and similarity can only decrease) as a result of a shift to $\alpha_{\delta,Qb} = 1$. For $c = con$ instead, when $\alpha_{\delta,Qb} = 1$ the term in Equation (31) has $\alpha_{P,Qb}$ optimally set to 1 so that $d\left(|\alpha_{P,Qb} - \alpha_{c,Qb}|\right) = 0 = d\left(|\alpha_{P,Qb} - \alpha_{\delta,Qb}|\right)$, as $\alpha_{c,Qb} = \alpha_{\delta,Qb} = 1$. So, it can only be smaller (and similarity can only be higher) than the original value of the term when $\alpha_{\delta,Qb} = 0$.

(ii) immediately follows by Proposition 6. ■

B.2 Bottom-up Contrast in Statistical Problems

Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [20] propose a model of selective attention to features of statistical problems driven by bottom-up contrast. Here we show that model is nested in ours, illustrating key predictions in judgments about iid draws and inference problems.

Beliefs about iid draws. A DM judges the likelihood that n draws of a fair coin produce a balanced sequence H_2 versus a full heads sequence H_1. [20] document that the prevalence of Gambler's fallacy, i.e. assigning a higher probability to the balanced sequence, is higher for longer sequences. Proposition 18 shows that similarity to the familiar inf category increases with sequence length, providing one mechanism for this finding. Here we show that bottom-up contrast of the share of heads also increases, which boosts similarity to inf and strengthens the effect. As in [20], i) the problem's event features are individual flips, and ii) the share of heads in a sequence, and contrast σ_i of event feature i depends on the probability difference across hypotheses when attention is paid to that feature alone, $\alpha_i = 1, \alpha_{j \neq i} = 0$. Among the problem's event features, only the contrast of the share of heads varies with n. In particular, using a special case of the contrast function of Equation (19), the contrast of share of heads $\sigma_{\mathcal{S}}$ for sequences with n flips is given by:

$$\sigma_{\mathcal{S}} = 1 + \frac{\left|\binom{n}{n/2} - 1\right| (0.5)^n}{\binom{n}{n/2}(0.5)^n + (0.5)^n}.$$

The bottom-up contrast of share of heads increases in n because homogeneous sequences become proportionately much rarer for larger n. By Proposition 6, when matching the problem to any category, the DM pays more attention to the share of heads for large n. It feels very striking, and thus attention grabbing, to obtain zero tails in 6 flips. By Proposition B.1, this promotes categorization in agnostic inference inf. As explained in the text, this ensures attention to share of heads which causes a "question substitution": the DM assesses the relative probabilities of a balanced versus a homogeneous sequence, as opposed to the specific sequences H_1 and H_2.

Inference. A computer chooses one of two urns, A and B, with base rates $\pi_A < 1/2$ and $\pi_B = 1 - \pi_A$, and picks a ball from it. The likelihood of taking a green ball from A is $q > 1/2$ and $1 - q$ from B. A DM evaluates the likelihood that a green ball comes from urn A. Bordalo et al. [20] document three facts: i) judgments are multi-modal, corresponding to neglect of either the base rate or the likelihood, ii) the prevalence of base rate neglect increases with the likelihood q, and iii) the prevalence of base rate neglect increases with Kahneman and Tversky's "Taxicab" reframing of the problem.
There are two event features, "urn selection" or base rate $\mathcal{U}$ and "share of green" $\mathcal{S}$. Fact i) follows from categorization into inference inf, triggering focus on the likelihood of green as well as base rate neglect, or into $freq$, which entails attention to the base rate and neglect of the likelihood. Relative to [20], our model explains why these specific profiles emerge, corresponding to two familiar categories of problems. Turning to fact ii), the bottom-up contrast of feature $\mathcal{S}$ increases in q, making categorization in inference inf more likely according to Proposition B.1. Lastly, the Taxicab frame increases the prominence of the signal accuracy feature; this feature is not prominent at all in the original frame, although it can be inferred from the base rates and the likelihoods. By drawing attention to it, the problem description promotes categorization into a $freq$ category for which the accuracy statistic is diagnostic. In fact, the accuracy statistic is framed as a frequency problem, namely "What is the probability that the witness says green when the car is green," for which the likelihood becomes the base rate. This leads the DM to answer this question instead of the original one, "What is the probability that the car is green if the witness says it is green", displaying full base rate neglect.

B.3 Lotteries

Here we show how the model of salience from payoff contrast in Bordalo, Gennaioli, and Shleifer ([21]) is nested in the current model, and examine well-known puzzles in choice under risk from the perspective of our model.
Consider the choice between monetary lotteries

$$X = (x_g, x_b; \pi) \ \ and \ \ W = (w_g, w_b; \beta),$$

which pay their upside prizes x_g and w_g with probabilities π and β, respectively, and pay x_b and w_b otherwise. X is a mean preserving spread of W, so $x_g > w_g$ and $\pi < \beta$. Each lottery has four features: its maximum and minimum payoffs (*hedonic*) and the maximum and minimum payoff events (*event*). Each lottery has two atoms of the form (u_g, u_b, e_g, e_b). The atom of a lottery state $s \in \{g, b\}$ specifies: i) in u_s the payoff in this state and in e_s its delivery event, and ii) the neutral values $u_{-s} = 0$ and $e_{-s} = \Omega$ for the features not corresponding to s.

Consider next how bottom-up salience drives attention in such problems. Given description prominence α_δ with $\alpha_{\delta,e_s} = 1$ for all $s \in \{g, b\}$ and quadratic distance d, Proposition 6 implies that, when matching to $c \in \{con, freq\}$, attention satisfies:[33]

$$\alpha_{P,u_s}(con) = \frac{\alpha + \alpha_{\delta,u_s}\sigma_{\delta,u_s}/F_{con}}{1 + \sigma_{\delta,u_s}/F_{con}}, \qquad \alpha_{P,e_s}(con) = \frac{\sigma_{\delta,e_s}/F_{con}}{1 + \sigma_{\delta,e_s}/F_{con}}, \tag{32}$$

$$\alpha_{P,u_s}(freq) = \frac{\alpha_{\delta,u_s}\sigma_{\delta,u_s}/F_{freq}}{1 + \sigma_{\delta,u_s}/F_{freq}}, \qquad \alpha_{P,e_s}(freq) = 1, \quad s \in \{g, b\}. \tag{33}$$

Ceteris paribus, compared to "frequency estimation", categorization in "consuming" boosts attention to payoffs but dampens it to probabilities, $\alpha_{P,u_s}(con) > \alpha_{P,u_s}(freq)$ and $\alpha_{P,e_s}(con) < \alpha_{P,e_s}(freq)$ for all $s \in \{g, b\}$.

Here we focus on the case, typical in lab experiments on choice under risk, where all payoff and event features are similarly prominently described. As in Bordalo, Gennaioli, and Shleifer [21], higher contrast of payoffs in state s (higher σ_{δ,u_s}) boosts attention to the payoff in this state. Higher event contrast (higher $\sigma_{\delta,e}$) boosts attention to probabilities. As in Proposition 6, the bottom-up forces are weaker in frequent categories (due to the terms $\sigma_{\delta,u_s}/F_c$ and $\sigma_{\delta,e}/F_c$).

At the equilibrium categorization, the DM's expected valuation gap for X over W is:

$$v(X) - v(W) = \Pr(con|P) \cdot \widetilde{v}(con) + \Pr(freq|P) \cdot \widetilde{v}(freq), \tag{34}$$

where the valuation gaps $\widetilde{v}(c)$ between X and W in each category c is computed using the above attention weights. We now consider different cases.

Common Ratio Effect. Suppose $x_b = w_b = 0$. It is well known that if $(100, 0.2; 0, 0.8) \sim (25, 0.8; 0, 0.2)$, then for many people $(100, 0.02; 0, 0.98) \succ (25, 0.08; 0, 0.92)$. This pattern violates expected utility, in which preferences are invariant to uniform scaling of probabilities. Our model

[33] To ease notation we set category contrast to $\sigma_{c,i} = 1$ for all $i \in \Phi$. Our qualitative results do not depend on this assumption.

explains this puzzle: as probabilities shrink, event contrast drops, triggering a shift towards a representation that focuses on payoffs, which benefits the riskier lottery.

Formally, the valuation gap $\widetilde{v}(c)$ between X and W in category c is given by:

$$\widetilde{v}(c) = \pi^{\alpha_{P,e_g}(c)} \cdot \alpha_{P,u_g}(c)\,(x_g - w_g) + \left[\pi^{\alpha_{P,e_g}(c)} - \beta^{\alpha_{P,e_g}(c)}\right] \cdot w_g \cdot \alpha_{P,u_g}(c). \qquad (35)$$

Since $\pi \cdot x_g = \beta \cdot w_g$, the right hand side increases in attention to payoffs $\alpha_{P,u_g}(c)$ and decreases in attention to probabilities $\alpha_{P,e_g}(c)$. Thus, categorization in *con* boosts risk taking while categorization in $freq$ boosts risk aversion.

In the problem above, the reduction of probabilities reduces event contrast, e.g., $|\pi - \beta|$ drops from $|0.8 - 0.2| = 0.6$ to $|0.08 - 0.02| = 0.06$. Payoff contrast is instead constant.[34] The drop in event contrast reduces $\alpha_{P,e}(c)$ in any category, but also boosts categorization into consuming. Overall, when differences in probabilities are "peanuts" the DM's attention is drawn away from them and towards the \$100 versus \$25 payoff difference, promoting risk taking.

The Fourfold Pattern and "Simplicity Equivalents". Suppose now that W pays a sure amount $\widetilde{w} = \pi \cdot x_g$. Hedonic features are whether the lottery pays more or less than $\widetilde{w}$, with corresponding event features. This is equivalent to the previous formalization of atoms provided that, for the sure thing, we set $w_g = w_b = \widetilde{w}$ and $\beta = \pi$. Payoff contrast σ_{δ,u_s} in state $s \in \{g, b\}$ increases in $|x_s - \widetilde{w}|$. The valuation gap for X over W in category c is now given by:[35]

$$\widetilde{v}(c) = x_g \cdot \left[\alpha_{P,u_g}(c) \cdot \pi^{\alpha_{P,e}(c)} \cdot (1-\pi) - \alpha_{P,u_b}(c) \cdot (1-\pi)^{\alpha_{P,e}(c)} \cdot \pi\right]. \qquad (36)$$

Equation (36) yields the so-called "fourfold pattern" in risky choice (Kahneman and Tversky [74]). Given that $x_g = \widetilde{w}/\pi$, upside payoff contrast increases in $|\widetilde{w}/\pi - \widetilde{w}| = \widetilde{w} \cdot \left(\frac{1-\pi}{\pi}\right)$, while downside contrast depends on $|\widetilde{w} - 0| = \widetilde{w}$. If the lottery is right skewed, $\pi < 0.5$, upside contrast is larger than downside contrast,

[34] We implicitly assume that contrast in state s is only computed for features that take a proper value in such state, and not using features that take values in a different state.

[35] The next formula is obtained by formalizing a safe lottery paying some amount w, when compared to a risky alternative (x_g, π) , as composed by two atoms $(w, 0, \{\omega : X(\omega) = x_g\}, \Omega)$ and $(0, w, \Omega, \{\omega : X(\omega) = x_b\})$. In words, it is modeled as a lottery with a good and a bad outcome that have respectively the probability π and $(1-\pi)$ of the alternative lotteries, but were both the good and bad outcome are equal to w. Qualitatively analogous results would be obtained by modeling the safe lottery as having equal good and bad outcomes with probability 0.5 that does not depend on the alternative under consideration. However, modeling it as a single atom would change the behavior described below.

and conversely if $\pi > 0.5$. Equations (32) and (33) then imply that in every category the DM focuses more on the upside if and only if $\pi < 0.5$. The DM is risk seeking for $\pi < 0.5$ and risk averse for $\pi > 0.5$, as in the fourfold pattern.[36]

Crucially, the same logic also explains Oprea's [98] evidence of similar behavior when choosing among similar riskless options. This arises because the similar description promotes the use of a common "consuming" representation in both domains. Consider a riskless choice where options are "riskless mirrors"of the lotteries above. Option A consists of 90 boxes with $x_g = \$2.5$ and 10 boxes with $x_b = \$0$, and option B consists of 100 boxes with $\widetilde{w} = \$2.25$. The subject is paid the total value of the chosen option divided by 100. The two options pay the same, but (when going down a multiple price list) many subjects exhibit a preference for B. Options have two features: the payoff in boxes $s \in \{g, b\}$ and their frequencies n_g and n_b. Also in this case, reasoning about payoffs prompts the "consuming" category, while reasoning about frequencies prompts the "frequency estimation" category.

When evaluating A versus B, the contrast between A's payoff in its g and b boxes with the payoff in B boxes prompts categorization as *con*. The DM focuses on evaluating payoffs of different boxes, reducing attention to their precise frequencies. From Equation (36), the valuation gap for A over B in category c is given by:

$$\widetilde{v}\,(c) = x_g \cdot \left[\alpha_{P,u_g}(c) \cdot \left(\frac{n_g}{100}\right)^{\alpha_{P,e}(c)} \cdot \left(\frac{n_b}{100}\right) - \alpha_{P,u_b}(c) \cdot \left(\frac{n_b}{100}\right)^{\alpha_{P,e}(c)} \cdot \left(\frac{n_g}{100}\right) \right]. \quad (37)$$

Equation (37) is equivalent to Equation (36) with risky lotteries, and so yields equivalent choice behavior. The fourfold pattern cannot come from preferences under risk, because the domains are different. It comes from the payoff evaluation representation, induced by high payoff contrast in both domains.

When choosing between riskless mirrors, the retrieval of the *con* category inhibits the correct adding up across boxes. We suspect that many subjects would be able to perform the addition if explicitly asked to do so, and would then be indifferent between A and B. Complexity is in representation, not in computation.[37]

[36] Relative to Bordalo, Gennaioli, and Shleifer [21], here payoff-contrast triggers a consuming representation, causing probability neglect ($\alpha_e = 0$ for $F_{con} \to \infty$). Thus, the fourfold pattern can also arise if the upside and downside payoffs are equally attended to, $\alpha_{P,u_g} = \alpha_{P,u_b}$, which is relevant for testing the model using attention data. We also predict that eliciting certainty equivalents should strengthen the fourfold pattern compared to, say, making binary choices or choosing a "probability equivalent π" to $\widetilde{w}$. Being in dollar units like payoffs, certainty equivalents are more similar to payoff evaluation, $c = con$.

[37] Riskless choice can also be contaminated by numerical tasks in other domains. In the left

B.4 Top-Down Contrast and Unstable Similarity

Tversky [122] famously showed that when people rate similarities between countries on a list, they judge Austria and Sweden as more similar when the list includes Hungary and Poland than when it includes Hungary and Norway. He explained this finding by the contrast principle. When Poland is on the list, political differences are contrasting, so Sweden and Austria are deemed similar. When Norway is on the list, geographic differences are contrasting, so Sweden and Austria are deemed dissimilar.

Here contrast arises top down: the only information people are given is country names, but these prompt focus on a feature contrasting among them (similarly to when seeing the "flight" label we think of the "crash" feature).

When assessing the similarity between Austria (a), Sweden (s), Hungary (h) and either Norway (n) or Poland (p), each atom $y \in Y$ lists the features of a country pair. The Austria-Sweden atom (a, s) reports two "hedonic" features: geographical distance $u_G(a, s)$ and political distance $u_L(a, s)$. Attention $(\alpha_{P,G}, \alpha_{P,L})$ to hedonics entails estimated distance

$v((a, s)(\alpha_{P,G}, \alpha_{P,L})) = \alpha_{P,G} \cdot u_G(a, s) + \alpha_{P,L} \cdot u_L(a, s) + (1 - \alpha_{P,G})\, \bar{u}_G + (1 - \alpha_{P,G})\, \bar{u}_L$,

which is used as an inverse measure of similarity. Because Austria and Sweden are intuitively more distant geographically than politically, $u_L(a, s) < u_G(a, s)$, they are judged more similar when attention to politics $\alpha_{P,L}$ is higher compared to geography $\alpha_{P,G}$.[38]

The DM has experienced two categories: problems of category G in which geographic features are learned or judged, and problems of category L in which political features are learned or judged. The former category attends to geography features while neglecting politics, $\alpha_{G,G} = 1 > \alpha_{G,L} = 0$, the latter does the reverse,

digit bias (e.g., List, Muir, Pope, and Sun [1]), consumers perceive \$9.99 to be dissimilar from \$10 despite the metric proximity. A number's digits are akin to event features in a sequence of i.i.d. draws. The fact that the left digit is usually more relevant prompts people to focus on this feature and neglect others, boosting the perceived difference between \$9.99 and \$10.

[38] Formally, these computations are performed under

$$\begin{aligned} u_G(a, s) &= u_G(a, n) = u_G(h, s) = u_G(h, n) = u_G(p, s) = u_G(p, n) = 1 \\ u_G(a, h) &= u_G(a, p) = u_G(p, h) = u_G(s, n) = 0 \\ u_L(a, h) &= u_L(a, p) = u_L(h, s) = u_L(h, n) = u_L(p, s) = u_L(p, n) = 1 \\ u_L(a, s) &= u_L(a, n) = u_L(n, s) = u_L(h, p) = 0. \end{aligned}$$

$$\alpha_{L,G} = 0 < \alpha_{L,L} = 1.$$

Top-down contrast in category G occurs only along geography, $\sigma(\kappa_G)$ where $\kappa_G = \{u_G(y)\}_{y \in Y}$ are the distances between the four countries. In category L, on the other hand, it only occurs along politics, $\sigma(\kappa_L)$, with κ_L accordingly defined. The description of the problem makes the country names fully prominent while it shrouds the hedonics.

Critically, top-down contrast of hedonic features changes with the described country names. When the DM is presented with a,s,h,n, variability along geography is high (n,s versus a,h) while variability along politics is low (n,s,a versus h), so $\sigma(\kappa_{G1}) > \sigma(\kappa_{L1})$, where 1 captures the list a,s,h,n. When the DM is presented with a,s,h,p, variability along geography is low (s versus a,h,p) while variability along politics is high (s,a versus h,p), so $\sigma(\kappa_{G2}) < \sigma(\kappa_{L2})$, where 2 refers to the list a,s,h,p.

Instability in similarity judgments arises because, by point i) in Proposition B.1, when the most contrasting category is G (country list 1), the DM retrieves this category and focuses on geography, holding a and s dissimilar. When instead the most contrasting category is L (country list 2), the DM retrieves this category and focuses on politics, holding a and s similar. The similarity judgment is unstable as documented by Tversky. As in the GF, upon seeing a, s, h, n the DM thinks "what a striking North-East separation between n,s and a,h!". This spontaneous association between the current task and geography does not just increase attention to this feature. It causes neglect of politics, which causes an unstable similarity judgment between a and s, shaped by irrelevant countries in the list.

CALL FOR NOMINATIONS

$300,000 Nemmers Prize in Economics

Northwestern University invites nominations for the Erwin Plein Nemmers Prize in Economics, to be awarded during the 2026–27 academic year. The prize pays the recipient $300,000. Recipients of the Nemmers Prize present lectures, participate in department seminars, and engage with Northwestern faculty and students in other scholarly activities.

Details about the prize and the nomination process can be found at **nemmers.northwestern.edu**. Candidacy for the Nemmers Prize is open to those with careers of outstanding achievement in their disciplines as demonstrated by major contributions to new knowledge or the development of significant new modes of analysis.Individuals of all nationalities and institutional affiliations are eligible except current or recent members of the Northwestern University faculty and past recipients of the Nemmers or Nobel Prize. **Nominations will be accepted until January 14, 2026.**

The Nemmers prizes are made possible by a generous gift to Northwestern University by the late Erwin Esser Nemmers and the late Frederic Esser Nemmers.

Northwestern

Nemmers Prizes • Office of the Provost • Northwestern University • Evanston, Illinois 60208
nemmers.northwestern.edu