A Cognitive Foundation for Perceiving Uncertainty[a]

J. Aislinn Bohren[b], Josh Hascher[c], Alex Imas[d], Michael Ungeheuer[e], and Martin Weber[f]

April 1, 2024

Abstract

We propose a framework where perceptions of uncertainty are driven by the interaction between cognitive constraints and the way that people learn about it—whether information is presented sequentially or simultaneously. People can learn about uncertainty by observing the distribution of outcomes all at once (e.g., seeing a stock return distribution) or sampling outcomes from the relevant distribution sequentially (e.g., experiencing a series of stock returns). Limited attention leads to the overweighting of unlikely but salient events—the dominant force when learning from simultaneous information—whereas imperfect recall leads to the underweighting of such events—the dominant force when learning sequentially. A series of studies show that, when learning from simultaneous information, people are overoptimistic about and are attracted to assets that mostly underperform, but sporadically exhibit large outperformance. However, they overwhelmingly select more consistently outperforming assets when learning the same information sequentially, and this is reflected in beliefs. The entire 40-percentage point preference reversal appears to be driven by limited attention and memory; manipulating these factors completely eliminates the effect of the learning environment on choices and beliefs, and can even reverse it. Our results have implication for the design of policy and the recovery of preferences from choice data.

Keywords: Choice Under Risk, Bounded Rationality, Perceptions of Uncertainty, Information, Beliefs, Attention, Memory, Description-Experience Gap

[a]Support from the Research Council of Finland is gratefully acknowledged. We are grateful to Peter Andre, Sam Hirshman, Lukas Mertes, Ryan Oprea, and Florian Zimmermann for valuable comments. Thanks to Haiying Lu and Kaushal Addanki for excellent research assistance.

[b]J. Aislinn Bohren: University of Pennsylvania, abohren@gmail.com.

[c]Josh Hascher: University of Chicago, josh.hascher@chicagobooth.edu.

[d]Alex Imas: University of Chicago, NBER, CESifo, HCEO. alex.imas@chicagobooth.edu.

[e]Michael Ungeheuer: Aalto University, michael.ungeheuer@aalto.fi.

[f]Martin Weber: University of Mannheim and the Centre for Economic Policy Research, London EC1V 3PZ, United Kingdom, weber@bank.BWL.uni-mannheim.de.

1 Introduction

People learn about the uncertainty involved in their choices through different ways. In some cases, they have access to information about all the potential outcomes and their likelihoods simultaneously. For example, price charts describe the historical returns of a stock and serve as a proxy for the underlying risk distribution; similarly, information about the potential outcomes and their odds are readily available for sports betting, casino gambling, and public lotteries. In other cases, people learn about the relevant distribution by sequentially sampling from it; the riskiness of a new business relationship or a stock is judged by observing and experiencing the relevant outcomes. In standard theory, the method through which one learns about uncertainty should not impact decision-making. As long as information is held constant, choices and beliefs about uncertainty should be the same regardless of whether one learns the information simultaneously or sequentially.

However, cognitive constraints may change people's perceptions of uncertainty as a function of the learning environment. We propose a framework that nests well-studied limits on attention and memory and derive predictions on how these limits affect beliefs and choices depending on whether information is presented simultaneously or sequentially.[1] In the case of learning from simultaneous information, memory plays little role and bottom-up channeled attention leads to the overweighting of unlikely but salient events. This can generate a preference for prospects that are most likely to underperform, but when they do outperform, they do so by a lot. A strict preference for such unlikely-outperformers can be interpreted as the overweighting of rare events. When learning through sequential sampling, however, imperfect memory recall implies that an infrequent event will be misremembered. This can shift preferences towards prospects that outperform most frequently—behavior that looks like the underweighting of rare events. Importantly, because the behavioral predictions are derived from basic cognitive processes, the learning environment is expected to change people's perceptions and mental representations of the risky prospects—learning from information presented simultaneously versus sequentially will lead to different *beliefs* about the relevant distributions.

The framework has implications for interpreting and predicting economic behavior. If the learn-

[1] As outlined in greater detail in Section 2, our framework builds on the models in the economics and psychology literature on limited attention, salience, and bounded memory (Bordalo et al., 2012; Li and Camerer, 2022; Bruce and Tsotsos, 2009; Anderson et al., 1998; Bordalo et al., 2022a).

ing environment shifts perceptions of uncertainty in a manner that systematically deviates from the objective distributions, then interpreting choices as reflecting underlying preferences becomes more difficult.[2] This underscores the need for both choice and belief data for identification of preferences and welfare analysis. The proposed cognitive foundation also implies that heterogeneity in learning environments, where some people learn from simultaneous information while others learn sequentially, will generate predictable heterogeneity in risk perceptions, beliefs, and behavior. For example, research in finance has documented that investors are attracted to lottery-like stocks and out-of-the-money options that outperform infrequently but mostly underperform (e.g., Barberis and Huang, 2008; Kumar, 2009; Boyer and Vorkink, 2014; Bryzgalova et al., 2023); in other contexts they are attracted to prospects that that mostly outperform but infrequently extremely underperform—a behavior commonly referred to as "picking up pennies in front of a steamroller" (e.g., buying yield enhancement products, Henderson and Pearson, 2011; Célérier and Vallée, 2017; Célérier et al., 2021; Vokata, 2021; Chesney et al., 2021). Our framework shows that such heterogeneity in choices is predicted by heterogeneity in how people learn about the associated uncertainty. Moreover, it provides a microfoundation for how heterogeneity in mental models can emerge *predictably* as a function of the learning environment (Kendall and Oprea, 2024; Handel and Schwartzstein, 2018).

To test the predictions of the framework, we presented people with choices between two risky assets that had the *same* underlying distribution of outcomes but differed in the outcome associated with each state. Outcomes were assigned to states such that one asset (Asset F) outperformed the other (Asset U) in all but one state, but Asset U outperformed Asset F by a lot in the final state. The distributions of outcomes for each asset are presented in Figure 1. People were given the same information about the relevant distributions. However, in some conditions they were presented with the state-by-state outcomes simultaneously, as in Figure 1, and in others they learned this information sequentially (one state at a time).[3] Here, a strict preference for Asset U can be interpreted as the overweighting of the unlikely state in which it significantly outperforms, while a

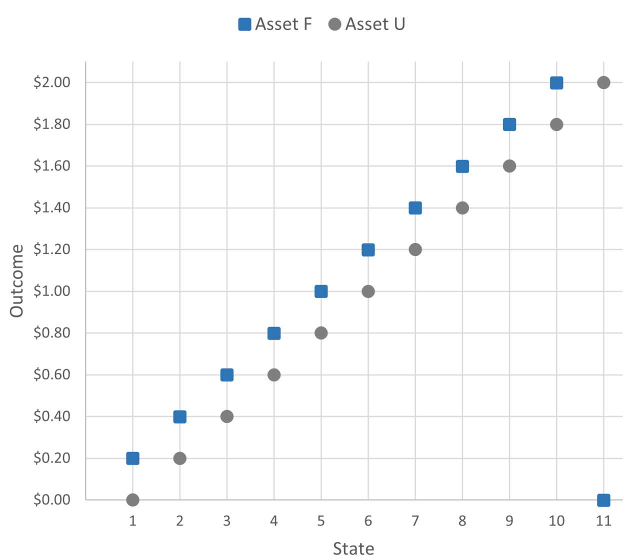
[2] Manski (2004) argues that data on behavior can be rationalized by different sets of preferences and beliefs. This implies that the recovery of preferences from choice data requires either strong assumptions (e.g., rational expectations) or belief data. Our findings show that the assumption of rational expectations is violated even in settings where people are given full information about the objective distribution, and that this violation varies *predictably* with the learning environment.

[3] To keep information constant, participants were sequentially presented with the outcome associated with each state without replacement, and were told that they would see the outcome associated with each potential state. This ensured that sampling error or differences in model uncertainty does not explain the results.

strict preference for Asset F can be interpreted as the underweighting of this state.

Our framework predicts that the salience of the state with the large performance difference will shift choices towards the unlikely-outperforming Asset U when the distributions are presented simultaneously. On the other hand, the chance of misremembering any given outcome will push people to select Asset F when they learn the same exact information sequentially. Indeed, Asset F was selected 42% of the time in the former environment and 85% of the time in the latter environment—a more than 40 percentage point gap in choices. These results are summarized in the Baseline conditions reported in Figure 2 below. Data on people's beliefs shows that these choices were a reflection of changes in people's perceptions of the prospects as a function of the learning environment: they had overoptimistic beliefs about the performance of Asset U (Asset F) when learning about the distributions simultaneously (sequentially). These results are particularly surprising because the underlying distributions of the assets were identical; cognitive constraints generated biased beliefs in both learning environments, but the direction of the bias reversed depending on whether information was presented simultaneously or sequentially.

Figure 1: Distributions of Outcomes: Frequently-outperforming Asset F versus unlikely-outperforming Asset U. All states are equally likely (1-in-11). The marginal outcome distributions of the two assets are exactly the same (discrete uniform on the $0.20 grid from $0.00 to $2.00), only shifted by one state.



Our proposed mechanism is inherently context-specific and comparative. A state is salient and thus overweighted based on *relative* differences in payoffs and visual factors within the choice set; a state is more or less likely to be misremembered depending on the number of other states and the similarity of elements within the choice set. Models where prospects are evaluated in isolation (e.g., prospect theory) would not predict a difference in choices. To test our framework directly, we exogenously manipulated context-specific factors that influence attention and memory, and studied the impact on choices and beliefs. In the simultaneous information environment, decreasing the salience of the unlikely state where Asset U outperformed shifted choices from 42% of people selecting Asset F ("Baseline") to 66% selecting Asset F ("No Salience Bias"), a preference reversal of 24 percentage points. Importantly, this manipulation was predicted to impact relative salience without affecting any underlying properties of the assets, including the state-specific outcomes.[4] The preference reversal against the unlikely-outperforming Asset U thus provides evidence for bottom-up attention as a driver for the overweighting of unlikely events in this and previous work (e.g., Kahneman and Tversky (1979)).

On the other hand, decreasing the scope for imperfect recall in the sequential information environment lowered the preference for Asset F by more than 20 percentage points.[5] Together, these two manipulations completely erased the gap in both choices *and* beliefs, which suggests that neither preference primitives for low-probability events (e.g., probability weighting) nor something inherent to learning from sequential information (e.g., reinforcement learning) is responsible for the gap.

Besides repetition, an important factor governing memory recall is feature-based similarity (Kahana, 2012; Anderson et al., 1998). When making a choice between two risky assets, people are more likely to recall outcomes that share similar features with an option when forming beliefs about the relevant distributions. However, if outcomes share similar features either across states or across assets, this creates the potential for *interference*, where a person attempting to recall the outcome of an asset in a given state may incorrectly recall an outcome from a different state or asset if the features are sufficiently similar. To test the interference mechanism directly, we included a third Asset D across two sequential information conditions: one where the third asset shared similar but payoff-irrelevant features with Asset U and another where it shared similar features with Asset F. The third asset's payoffs were similar across conditions and generally better than those of the others. Importantly, people could not choose this third asset in either condition: the choice was

[4] As discussed further in Section 3, we validated our assumption on shifting salience using Graph Based Saliency tools from cognitive psychology.

[5] We did this in two ways: having people keep a record of the sample ("No Recall Bias"), and reducing the length of the sample to minimize memory requirements ("Minimal Requirement"). Both generated a similar reduction in the selection of Asset F.

always between Asset U and F as in the Baseline study.

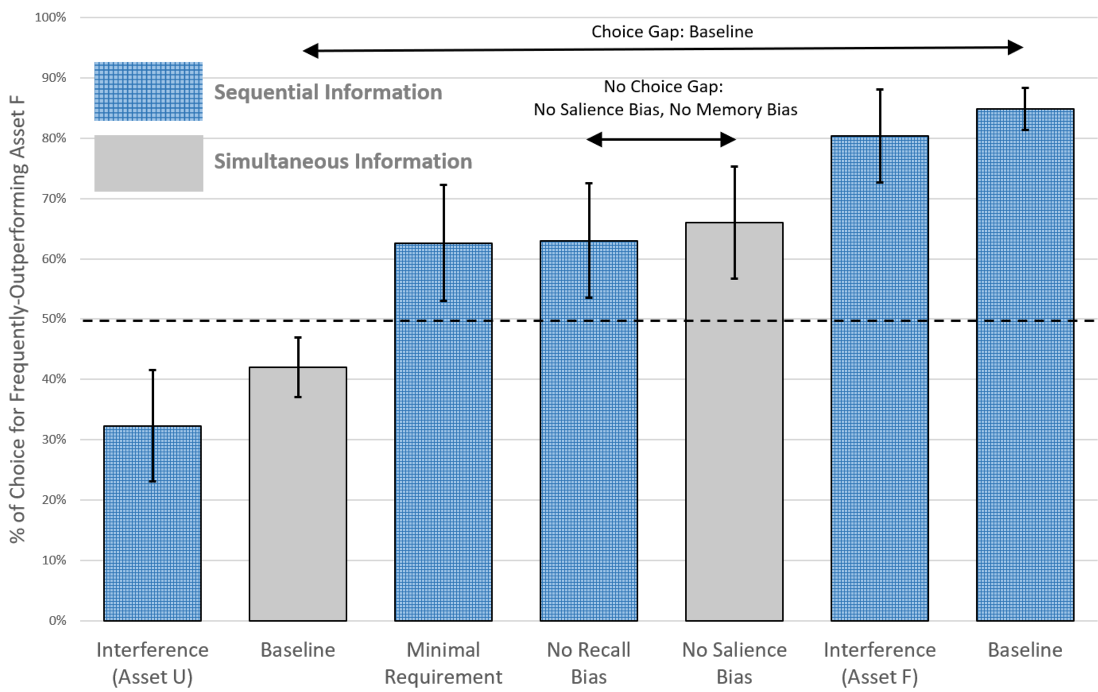
Our framework predicts that despite the choice and learning environment remaining the same, people will be more likely to choose Asset U when the third asset is similar to it—because they incorrectly recall the outcomes of Asset D when forming beliefs about Asset U—than when the third asset is similar to Asset F. The results were stark: despite the choice and learning environment being held constant, people went from selecting the unlikely-outperforming Asset U less than 20% of the time when the third asset was similar to Asset F ("Interference (Asset F)") to selecting Asset U more than 60% of the time when the third asset was similar to Asset U ("Interference (Asset U)").[6] This more-than-40 percentage points shift in choices in the sequential information environment led to an even higher (revealed) preference for the unlikely-outperforming asset than the simultaneous information environment. People's beliefs matched the choice patterns: they were more optimistic about the performance of Asset U (F) when the third asset was more similar to Asset U (F).

Figure 2 summarizes results across conditions. Despite facing the same decision and having the same information, the interaction between the learning environment, attention, and memory shifted people's choices by nearly 50 percentage points. This dramatic difference in revealed preferences highlights the importance of understanding the role of cognition in driving choice behavior. The analogous data on beliefs shows that the standard assumption that people's perception of risk matches the objective distributions conveyed to them may not hold. Exploring the factors which affect mental representations is critical for understanding risky choice.

Our findings have implications for the observed "decision-experience" gap in risky choice (see Hertwig et al. (2004), and de Palma et al. (2014) for review). In prior work documenting this gap, people select between a safe option and a risky gamble that has a large but unlikely payout. In the description environment, information about the outcomes and their respective likelihoods is presented simultaneously. In the experience environment, people learn this information by sequentially sampling from the distribution of the gamble, with replacement; they can stop sampling whenever they want and then proceed to make a choice. People were more likely to select the risky asset in the former environment than the latter. Several explanations have been proposed

[6]The shift in choices resembles a "decoy effect" Tversky and Simonson (1993) but is conceptually different because the manipulated attributes were payoff-irrelevant.

Figure 2: Learning about Uncertainty. Our Baseline Experiment documents a highly significant choice gap as a function of the learning environment: Most choose the unlikely-outperforming Asset U when learning from simultaneous information but switch to choosing the frequently-outperforming Asset F when learning from sequential information, despite both assets having the same marginal distributions. This gap can be eliminated and reversed by manipulating memory and attention. Having participants keep a record of the sample ("No Recall Bias"), reducing the length of the sample to minimize memory requirements ("Minimal Requirement"), and interfering with memory formation by introducing an attractive third asset that is similar to Asset U ("Interference (Asset U)") reduces the fraction of choices for Asset F after sequential information from 84% down to 32%. Meanwhile, minimizing the salience of the infrequent, extreme state ("No Salience Bias") increases the fraction of choices for Asset F after simultaneous information from 42% to 66%.



for this shift in choices. In the case of learning from description, the higher preference for the risky asset has been attributed to probability weighting, which leads to the overweighting of unlikely events (Kahneman and Tversky, 1979). In the model, objective probabilities are transformed through a probability-weighting function—typically assumed to be a primitive—which leads people to overweight low probability outcomes and underweight high probability outcomes (Tversky and Kahneman, 1992). Some have argued that the apparent underweighting of unlikely events when learning from experience is due to differences in information. For example, Fox and Hadar (2006) propose that people are more likely to go for the safe option because they do not even know about the unlikely payout of the gamble—they simply may not sample enough to have the same information about the distribution as those in the experience treatment (see Cubitt et al. (2022) for a similar conjecture). Others argue that the gap is driven by reinforcement learning (Erev and Barron, 2005) or simple tallying heuristics (Hills and Hertwig, 2010). Notably, Erev et al. (2010) do propose limited memory as one of many potential explanations, but do not test this mechanism directly.

We derive the "description-experience" gap from first principles—namely, bounded attention and memory—and present direct evidence for these mechanisms within the same paradigm. Importantly, we predict and show that the gap in choices is driven by a difference in perceptions of and beliefs about uncertainty. We do so in a setting that is explicitly designed to rule out the explanations that have been previously proposed in the literature. The experiment 1) involves a choice between assets with the same outcome distribution (which rules out primitives such as probability weighting), 2) ensures that people have the same information about the distribution across learning environments (which rules out sampling error), 3) abstracts from differences in hedonic feedback before the choice is made (which rules out reinforcement learning), and 4) allows for the explicit manipulation of factors affecting attention and memory (which allows for identification of the mechanism). The fact that we can eliminate and even reverse the "description-experience" gap by manipulating the relevant cognitive constraints suggests that the gap is not driven by something inherent to learning from description or experience per se.

Our results contribute to the literature on experience effects in consumer choice and financial decision making. Studies have documented long-lasting effects of personal experiences on economic decision-making; for example, Malmendier and Nagel (2011) find that negative individual stock-market experiences lead to more conservative investing decades later (see Malmendier and Wachter (2021) for review). This literature examines the influence of an individual's personal sample compared to the full historical sample, finding that the former is overweighted in both beliefs and decisions.[7] We show that people can either over- or underweight unlikely 'extreme' states depending on the learning environment. Our findings suggest that while 'extreme' events are underweighted when learning from sequential information alone (e.g., experiencing observations throughout one's life), factors that are important for memory recall such as cue-based similarity

[7]The personal experience literature does not examine the weighting of unlikely "extreme" states compared to more likely "moderate" states within the individual's sample, which is the focus of our study.

(e.g., media coverage of similar events) or interference can lead to them to be overweighted. Another strand of literature shows that generating *artificial* experience through sequential sampling can be used to affect beliefs about returns and investment choices (e.g., Kaufmann et al., 2013; Laudenbach et al., 2022). We contribute to this research on natural and artificial experience effects by highlighting the interaction between bounded cognition and how information is presented as the essential link between beliefs and choices in sequential learning environments.

We also contribute to the more conventional portfolio choice and asset pricing literature by varying dependence while carefully holding marginal distributions constant. Dependence is intrinsic to models of optimal portfolio choice and capital market equilibrium going back to Markowitz (1952) and Sharpe (1964). In these models, dependence between asset returns is important because it results in non-diversifiable risk (systematic risk or 'beta'), which makes portfolios more risky for individuals and depresses prices in the aggregate. Going beyond the effect of dependence on portfolio risk, we provide the underpinnings of recent studies focusing on effects of dependence on beliefs, portfolio choice, and asset prices when investors learn from realistic samples of returns and prices (Ungeheuer and Weber, 2021, 2023). Incorporating dependence is unusual in behavioral finance, where typical approaches assume narrow framing when investors evaluate individual assets (e.g. Barberis et al., 2016, where investors evaluate stocks in isolation using the Cumulative Prospect Theory value function and a sample of 60 monthly returns). The few exceptions that explicitly consider dependence being Cosemans and Frehen (2021), who correlate historical stock returns with return predictors based on Salience Theory, and Barberis and Jin (2023) who consider the implications of reinforcement learning for investor behavior. We contribute to this literature by showing how dependence can affect capital markets through a fundamentally new interaction between learning environments, memory and attention mechanisms (not related to portfolio risk), and belief formation.

Finally, our findings inform the line of work exploring how cognitive constraints reconcile and rationalize seemingly anomalous economic phenomena in a variety of domains. Models of bottom-up attention and salience are used to explain the Allais paradox and risk-seeking behavior (Bordalo et al., 2012), insensitivity to taxes (Chetty et al., 2009; Taubinsky and Rees-Jones, 2018), and behavior in strategic interactions (Li and Camerer, 2022); see Bordalo et al. (2022b) for a review. Models of bounded memory can explain projection bias, context-dependent consumer choice, availability and representativeness heuristics (Bordalo et al., 2022a, 2020), as well as overreaction to information (Enke et al., 2020; Hartzmark et al., 2021), and belief-decay (Graeber et al., 2023). Recent studies build on evidence from neuroscience to model how the difficulty of perceiving and adapting to outliers can explain the neglect of outlier risk (d'Acremont and Bossaerts, 2016; Payzan-LeNestour and Woodford, 2022). Models of constructed preferences are used to rationalize anchoring effect and provide interpretations for seemingly-coherent demand curves. Finally, models of cognitive noise and complexity are used to explain small-stakes risk aversion (Khaw et al., 2021), insensitivity to decision-relevant parameters (Enke and Graeber, 2023), the four-fold pattern of risk (Oprea, 2022), and non-constant time discounting (Enke et al., 2023). We contribute to this literature by identifying an important interaction between cognitive constraints and the learning environment, and documenting its effect on perceptions of uncertainty and choice behavior.

The rest of the paper proceeds as follows. In Section 2, we outline the framework and predictions on uncertainty perception and risk-taking as a function of the learning environment. We introduce our experimental design in Section 3. Section 4 presents the results of our experiments for both choices and beliefs, and discusses alternative explanations. We conclude in Section 5 by discussing potential implications of our findings.

2 Cognitive Constraints and the Learning Environment

We now discuss a framework of bounded memory and attention and present predictions on how these mechanisms interact with the learning environment to impact risky choice. Appendix A derives the predictions formally. This framework is tightly linked to our experimental design presented in Section 3, which was developed to allow for theory-driven manipulations of the relevant cognitive constraints.

In the framework, the decision-maker (DM) first recalls information about an asset's distribution of outcomes and then evaluates this information when making a choice. With perfect memory and no attentional distortions, the learning environment has no impact on choice—recalled information at the time of decision is the same regardless of whether outcomes are sampled sequentially or the distribution is provided simultaneously—and the DM makes a decision in line with Expected Utility Theory. However, allowing for imperfect recall can lead to behavior that seems to underweight unlikely outcomes, while attentional distortions will lead to the overweighting of outcomes in unlikely but salient states, conditional on being recalled.

A DM evaluates a set of N assets. Let A^i denote asset i for $i = 1, ..., N$ and $\omega \in \Omega$ denote each possible state, which occurs with probability π_ω. Let $x^i_\omega \in \mathbb{R}$ correspond to the outcome of A^i in state ω. The DM's subjective value from a given outcome is denoted $v(x^i_\omega)$. The DM with perfect memory and no attentional distortions will compute the following value $V^R(A^i)$ for each asset:

$$V^R(A^i) = \sum_{\omega \in \Omega} \pi_\omega * v(x^i_\omega). \tag{1}$$

where the superscript R corresponds to the rational benchmark. A DM in our setting will depart from this baseline due to a) imperfect recall of outcomes because of memory constraints and b) over/underweighting likelihoods due to bottom-up attention. To preview the results, recall will be adversely affected by the number of states as well as the similarity of outcomes across assets; over/underweighting will be affected by the salience of states, which is a function of both the payout distribution and visual factors.

2.1 Gathering information

To make a choice, the DM assembles information about the distribution of risk either from being presented with it directly or by learning through sequential sampling, and uses it to form a mental representation of the mapping between states and outcomes. In the case of the sequential sampling, at the time of choice the DM must accrue this information through memory recall. The process of memory recall in our framework builds on the models of Bordalo et al. (2020) and Anderson et al. (1998). Let $S(u, v) : E \times E \to [0, 1]$ measure the similarity between outcomes u and v in a DM's set of sampled outcomes E, where $S(u, v) = 1 \iff u = v$. Let χ^i_ω be the remembered outcome in state ω for asset A^i, $s^{ij}_\omega = S(x^i_\omega, x^j_\omega)$ be the similarity between the outcomes in state ω for assets i and j, and $\mathbf{s}^i_\omega$ be the vector of similarities between asset i and $j \neq i$ in state ω. The DM correctly recalls a given outcome (i.e., $\chi^i_\omega = x^i_\omega$) with probability $r(\mathbf{s}^i_\omega, |\Omega|) \in [0, 1]$. Assume r is decreasing in each similarity s^{ij}_ω and decreasing in the cardinality of the state space $|\Omega|$.[8]

With probability $1 - r(\mathbf{s}^i_\omega, |\Omega|)$, the DM instead recalls the average between the target outcome

[8]For example, a simple recall function in the spirit of Bordalo et al. (2021) would be $r(\mathbf{s}^i_\omega, |\Omega|) = \frac{1}{|\Omega|\left(1+\sum_{j\neq i} s^{ij}_\omega\right)}$.

and the other most similar outcome in that state. That is, the DM experiences two frictions in her ability to recall an asset's outcome in a given state: the similarity of outcomes across assets for a given state—the standard *interference* channel in memory research—and the number of states she potentially needs to consider. Correct recall increases as outcomes grow more dissimilar across assets—which weakens interference—and the number of states decreases.

2.2 Forming mental representation

Once the DM has a mental representation of the mapping between states and outcomes, either through the recall process described above or by being presented with it directly, she forms a mental representation of the respective likelihoods of each outcome. This stage of the mental representation is subject to attentional distortions as a function of state-specific salience. Namely, states that are more (less) salient will attract more (less) attention and be over(under)weighted in the decision-making process.

A large literature in economics and psychology has shown the importance of attention in risky choice. Limited attention is channeled towards some states over others as a function of their salience (Brandstätter and Körner, 2014; Spitmaan et al., 2019; Bordalo et al., 2022b). As a result of channeled attention, states that are more salient are overweighted in the decision-making process compared to states that are less salient (Frydman and Mormann, 2016). In the context of risky choice, a state's salience is determined by both visual stimuli (Frydman and Mormann, 2016; Bose et al., 2022) and payoffs (Bordalo et al., 2013).[9] Notably, as shown in Bose et al. (2022), Li and Camerer (2022), and Bruce and Tsotsos (2009), amongst others, the salience of an object can be predicted using algorithms such as the Graph-Based Visual Saliency model (Harel et al., 2006). We use this technique to validate our assumptions on the salience ranking of states in our experimental studies.

When forming a mental representation, the weight assigned to a state is a function of its salience. Specifically, in the framework, states are first are ranked by their relative salience. Then, objective state probabilities are replaced with distorted probabilities, where more salient states are overweighted and less salient states are underweighted. Let $\sigma(\chi_\omega^1, ..., \chi_\omega^N)$ be a function mapping

[9] For empirical evidence on the importance of state-based salience in the domain of belief-updating, see Ba et al. (2022).

recalled outcomes to state salience, and k_s be the relative ranking of state ω by salience.[10] The weight then given to state ω is:

$$w_\omega = \frac{\delta^{k_\omega}}{\sum_{\omega' \in \Omega} \delta^{k_{\omega'}} * \pi_{\omega'}}, \tag{2}$$

where $\delta \in (0, 1]$, with $\delta = 1$ corresponding to no salience-weighting and $\delta \approx 0$ corresponding to extreme weighting that assigns all mass to the most salient state. Finally, the DM evaluates each asset by computing:

$$V(A^i) = \sum_{\omega \in \Omega} \hat{\pi}_\omega * v(\chi^i_\omega), \tag{3}$$

where $\hat{\pi}_s = \pi_s * \omega_s$.

2.3 Choice behavior

Consider a decision environment with a choice between two assets A^U and A^F and a finite number of equally-likely states $\omega \in \Omega$, where $x^F_\omega > x^U_\omega$ for all but one state ω^*, and $x^F_{\omega^*} < x^U_{\omega^*}$. Finally, as in our experimental setting, for each outcome x^U_ω in A^U, there exists a state ω' where A^F has an equivalent outcome, $x^F_{\omega'} = x^U_\omega$ and vice-versa. A key property of this environment is that $V^R(A^U) = V^R(A^F)$. Our first prediction is on the choice between assets under perfect recall.

Prediction 1. *With no memory constraints, a DM with salience-distorted attention will prefer A^U to A^F.*

This prediction is relevant to the "simultaneous" learning environment where the DM is presented with information about the potential outcomes simultaneously. There, imperfect recall does not play a role since all of the information necessary to form an accurate mental representation is present at the time of choice. Our second prediction introduces the possibility of imperfect recall of the relevant outcomes.

Prediction 2. *Consider a setting with imperfect recall. Suppose the similarity between outcomes is nonzero across both states and assets. Then there exists a number of states $|\Omega^*|$, such that for $|\Omega| > |\Omega^*|$, the DM will prefer A^F to A^U.*

The intuition for the second prediction is straightforward. If the DM misremembers the state

[10] See Appendix A for the specific properties of σ.

in which $x_\omega^U > x_\omega^F$ while remembering any number of states where $x_\omega^U < x_\omega^F$, then she will prefer A^F to A^U. Consider the structure of A^U and A^F, where the former outperforms the latter in one state while the latter outperforms the former in each of the others. This implies that as the number of states grows, the chance that the DM will misremember the state in which $x_\omega^U > x_\omega^F$ while remembering a state where $x_\omega^U < x_\omega^F$ will increase as well. Since imperfect recall plays a role in the "sequential" learning environment but not the "simultaneous" one, taken together, the first two predictions imply a preference reversal: A^F will be selected over A^U in the former and A^F will be selected over A^U in the latter.

Our third prediction is on how similarity impacts choice behavior through the recall process.

Prediction 3. *Consider a setting with imperfect recall and fix the number of states $|\Omega|$. The likelihood that A^F is preferred to A^U is increasing in the similarity of outcomes both across assets within a state and across states within an asset.*

The intuition for the third prediction follows from the second. Similarity between outcomes increases the chance that they will be misremembered. The relative impact of similarity is higher for the outcomes of infrequent states simply because there are less of them. Thus, conditional on a number of states $|\Omega|$, the DM is (weakly) more likely to prefer A^F to A^U as the similarity between outcomes increases.

We build on this intuition for our last prediction, which directly tests the interference aspect of imperfect recall. Consider a third asset A^D that is better than both A^U and A^F, which the DM learns about but cannot choose. Interference predicts that one could reverse the preference between assets in the "sequential" environment when the third asset is more similar to asset A^U than asset A^F.

Prediction 4. *Suppose a third asset A^D is included, which dominates both A^U and A^F in every state. The likelihood that A^U is preferred to A^F is increasing in the similarity between the outcomes of A^D and A^U.*

The logic for this prediction comes from the property of interference that underlies most models of memory. When an outcome from a state is misremembered, the outcome that replaces it is more likely to be associated with a similar source. If the "replacement" outcomes from the similar source are superior to those of both assets in the choice set, then misremembering an outcome from

A^U will increase its attractiveness. Including the similar asset also increases the chances that an outcome from A^U is misremembered in the first place. Both forces lead to an increased preference for A^U in the sequential learning environment.

Notably, as we show in Appendix A, including a third asset A^D that is more similar to A^F maintains the preference for A^F over A^U in the sequential learning environment.

3 Experimental Design

In this section, we first describe the essential features of our Baseline paradigm that is used to study the impact of the learning environment on risky choice.[11] Participants ($N = 785$) were recruited using the Prolific crowdsourcing platform (Gupta et al., 2021) and restricted to investors from the US and UK. Summary statistics for participants are reported in Table IA1 in the Internet Appendix.

The experiment's overall structure is illustrated in Figure IA1. Participants first received introductory instructions. These included descriptions of the upcoming tasks, the information they would have access to, as well as an explanation of how they would be paid. Participants then went through two rounds of investment decisions. In each, they saw information about two risky assets. As outlined below, participants learned about each asset's distribution by either being presented with the outcomes simultaneously or by sequentially sampling them. After being given this information, participants chose between the two assets and stated their beliefs about the assets' outcomes. Beliefs about risk were elicited in the form of probabilities for an outcome less than \$1, equal to \$1, or greater than \$1. Both assets' actual distributions' were discrete uniform between \$0 and \$2, so that \$1 is always the mean and median.[12] For the choice between assets, we randomize the order of presentation both within-participants across rounds and across participants. To account for potential spill-over effects, half of participants were first asked for their choice and incentivized accordingly ("choice group"), and the other half was first asked for their beliefs and incentivized for accuracy ("beliefs group"). The experiment concluded after the second round with

[11] The experiment was pre-registered on AsPredicted.org, see this link. Complete instructions are reported in a supplementary document provided under this link.

[12] Danz et al. (2022) show that the binarized scoring rule results in bias in elicited beliefs and higher error rates. They argue that incentivizing truthful reporting using belief quantiles, which we use here, results in more truthful reporting.

some final questions (e.g., self-assessed risk aversion) and participants were informed about their bonus payment, which depended on their incentivized task. Across all treatments, we also elicited participants' strength of preference, which included the option to indicate that the choice between assets corresponded to indifference.

One of the two rounds was randomly selected for the bonus payment at the end of the experiment. The "choice group" was later paid with a bonus drawn from the selected asset's distribution (outcome between \$0 and \$2, on average \$1). The "beliefs group" was paid for a randomly selected belief (e.g., the average of Asset F's outcomes) with a bonus of \$2 if their stated belief is within $+/-0.03$ of the statistically correct value. Including a base fee, the average hourly salary for our Baseline Experiment was more than \$13.

3.1 Varying the Learning Environment

The joint distribution of the two risky assets is described in Figure 1. Both assets exhibited precisely the same marginal outcomes, uniformly distributed between \$0 and \$2.[13] However, the structure of dependence is asymmetric. Asset F moderately outperformed Asset U most of the time, while underperforming by a large margin once.[14] In the Baseline study, participant made choices between assets with a "short" distribution consisting of 11 states (or "situations") in one round, as shown in Figure 1, as well as a "long" distribution with 51 states in another round. The order of the two rounds was counterbalanced, i.e., some participants first saw the short and then the long distribution, or vice versa.

Participants learned about the joint distribution of the assets in two different environments. These environments varied between-subject and represent our main treatments. Half the participants were given information that described the outcomes of both assets in all states (or "situations"), along with the probability of each state. This constituted our Simultaneous Information treatment. The other half learned about the two risky assets by sequentially sampling each potential state and seeing the outcomes from the joint distribution. Each was told that 1) they would see the outcomes of each potential state and 2) that each state was equally likely to be realized. This constituted our Sequential Information treatment. It is important to note that presenting participants with

[13]This design was inspired by the paradigm used in Paterson and Diekmann (1988) to study violations of Expected Utility axioms.

[14]The participants' instructions referred to Asset U and F as Asset A and B, respectively.

every state in the Sequential Information treatment ensured that they were exposed to the same information about the relevant distributions as those in the Simultaneous Information treatment at the time of choice.

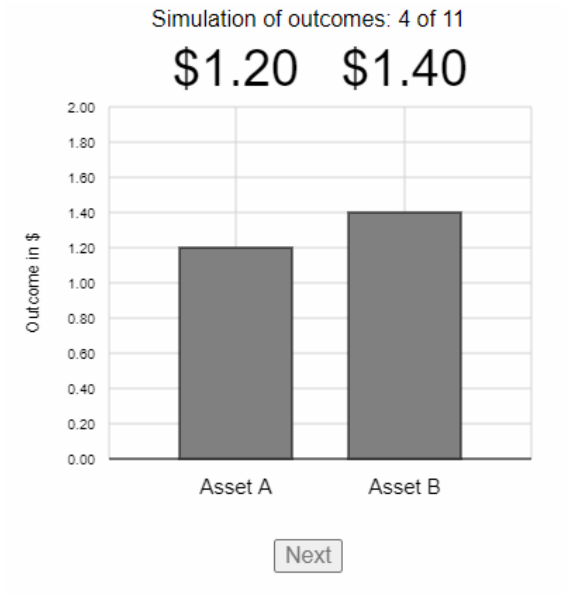
Figure 1 illustrates the Simultaneous Information learning environment. The graph shows the outcome of both assets in all states. In the instructions before the first round, participants also received an explanation of the graph and the information that all states are equally likely. They needed to pass a comprehension test to continue to the first round. Underneath the graph, a verbal statement also describes the distribution, as is common in traditional decision-making experiments. The statement mentions all joint outcomes and their probabilities, thus fully describing the joint outcome distribution.[15] In the Simultaneous Information treatment, participants could also view the graph when making the choice between assets and reporting beliefs, removing the need to memorize the distribution.

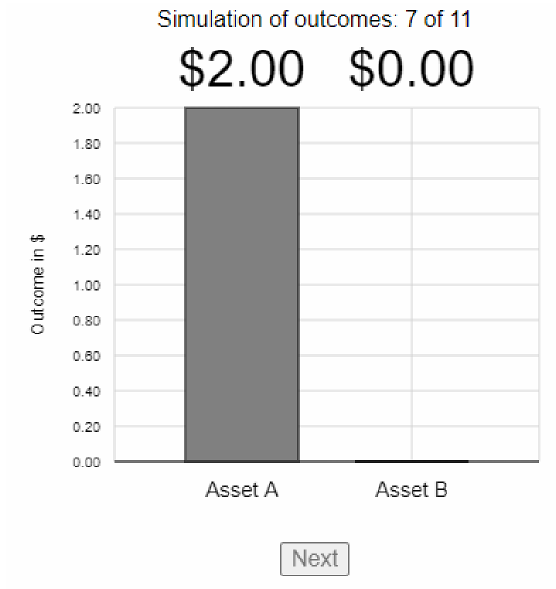
Figure 3 illustrates the Sequential Information learning environment, which uses both graphical (bar chart) and numerical information (realizations in $) to clearly present all outcome pairs.[16] Participants drew 11 (51) outcome pairs in the short (long) round of the experiment. Each sample is fully representative—presenting participants with the outcomes in every state—so that there is no sampling error. Additionally, each sample is randomly ordered, so that particular outcome orders cannot explain systematic treatment effects. Random ordering also allows us to analyze the importance of recency and primacy effects, which we explore in Section D.1.4. Participants chose how long to view the outcomes of each state and when to move on by clicking "Next." There was no option to go back. We collected viewing times for each state, which we use to analyze attention effects in Section 4.3. In the Sequential Information treatment, participants could not view a "simultaneous" summary of the relevant distributions while making their asset choices and reporting their beliefs.

[15] See Figure IA2 for the full Simultaneous Information environment including the verbal statement. The graph was designed to highlight the differences between assets using both visual presentation and text. Follow up treatments which remove both the text and box from the graph, while keeping the state ordering constant, yielded the same results as this Baseline Simultaneous Information treatment.

[16] This sampling design was inspired by the experiment in Ungeheuer and Weber (2023) who study the impact of outperformance on investment choice.

Figure 3: Baseline Experiment – Sequential Information Treatment.





3.2 Discussion of Design

There are several noteworthy aspects of the Baseline experimental design. First, showing a a fully representative sample in the Sequential Information learning environment ensures that participants have been exposed to the same information at the time of choice, which accounts for the oft-cited sampling error explanation for the description-experience choice gap (Fox and Hadar, 2006; Cubitt et al., 2022). Second, presenting the state-by-state outcomes in the Simultaneous Information learning environment accounts for the "mere-presentation effect" as a potential explanation (Erev et al., 2008).[17] The same design choice also accounts for the "unpacking-and-repacking" explanation (de Palma et al., 2014) which also relies on a difference in emphasis on unlikely outcomes between the Simultaneous and Sequential Information learning environments. Third, we used assets with the same underlying distributions to rule out explanations based on primitives such as probability weighting, since models like original or cumulative prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) would not predict a preference for either asset in the Simultaneous Information environment. Finding a strict preference for the unlikely-outperforming Asset U in the

[17] Erev et al. (2008) argue that the description-experience choice gap is due to the typical presentation of lotteries in the description learning environment, e.g., $((\$x, p; \$0, 1-p) = (\$8, 0.1; \$0, 0.9)$, affording propositional representation to the unlikely events. The same lottery's presentation through sampling is afforded analogical representation, where the unlikely outcome is also unlikely to be presented. Our design accounts for this by presenting the outcome of each state individually both in the simultaneous and Sequential Information learning environments.

Simultaneous Information environment would thus provide evidence for bottom-up attention as a driver for the overweighting of unlikely events. Finally, employing a sequential learning environment where outcomes were not payoff relevant accounts for reinforcement learning as an explanation for any observed differences in choices (Erev and Barron, 2005).

Our design differs from prior work on risky choice by including a graphical depiction of state-by-state distributions of outcomes. We did this for several reasons. First, the graphical depiction ensures that the presentation of events is proportional to their probability, as is the case in the Sequential Information learning environment. Second, as outlined further below, the visual representation allows us to manipulate multiple aspects of salience exogenously and to then validate those manipulations using existing tools from cognitive psychology.

3.3 Testing the Mechanism

We designed five additional treatments to directly test the proposed framework by exogenously manipulating memory and attention. All variations were theory-driven, motivated by the model outlined in Section 2 and Appendix A. The framework predicts that manipulating salience can drive choices away from the unlikely-outperforming Asset U towards the frequently-outperforming Asset F in the Simultaneous Information environment. It also predicts that manipulating memory constraints can decrease the preference for the frequently-outperforming asset in the Sequential Information environment, and reverse the preference altogether.

3.3.1 Manipulating Salience

We sought to manipulate salience without changing the state-by-state distribution of the assets. To do so, we leveraged the Attention based on Information Maximization (AIM) model of Bruce and Tsotsos (2009) to change the presentation of the states from one where the salient state clearly deviated from the local environment—and thus attracted attention—to a depiction where this deviation was less unexpected.[18] We did this by changing the order of the states from the predictably-increasing pattern in the Baseline treatment—where differences in outcomes in adjacent

[18] In the AIM model, salience emerges from first principles of bounded rationality—efficient coding and information theory. The salience of some component (in our case, the state) is a function of the local information relative to the surrounding information; a state is more salient if the content is unexpected given the surrounding context. See Appendix A for further discussion.

states are relatively small except for where Asset U outperforms—to a randomized order where adjacent states could have large differences in the levels of outcomes—where a large outcome difference within a state would be less surprising. The information presented to participants ($N = 100$) in this No Salience Bias treatment is depicted in Figure 4 below; otherwise the design was identical to the Baseline Simultaneous Information treatment.[19] It is important to note that the underlying distributions and the numbers associated with each state were the same across the Baseline and No Salience Bias treatments.

To test our assumption that the No Salience Bias treatment decreased the salience of the state where the unlikely-outperforming asset was superior (state 11), we used the bottom-up Graph-Based Visual Saliency model and the associated algorithm proposed in Harel et al. (2006).[20] The output of the model for both treatments is presented in the bottom row of Figure 4. As predicted, in the Baseline Simultaneous Information treatment, the outcomes associated with the state where Asset U outperforms are by far the most salient; the large contrast between payoffs relative to the local environment is predicted to channel more attention to that state than other states. However, changing the presentation such that the differences in payoffs within a state are less surprising substantially decreased the salience of state 11. No state is clearly dominant in terms of salience in the No Salience Bias treatment, implying that salience is less likely to bias choice towards a particular asset in this setting. Based on Prediction 1, people should choose Asset U less in the Random Order treatment than in the Baseline Simultaneous Information treatment.

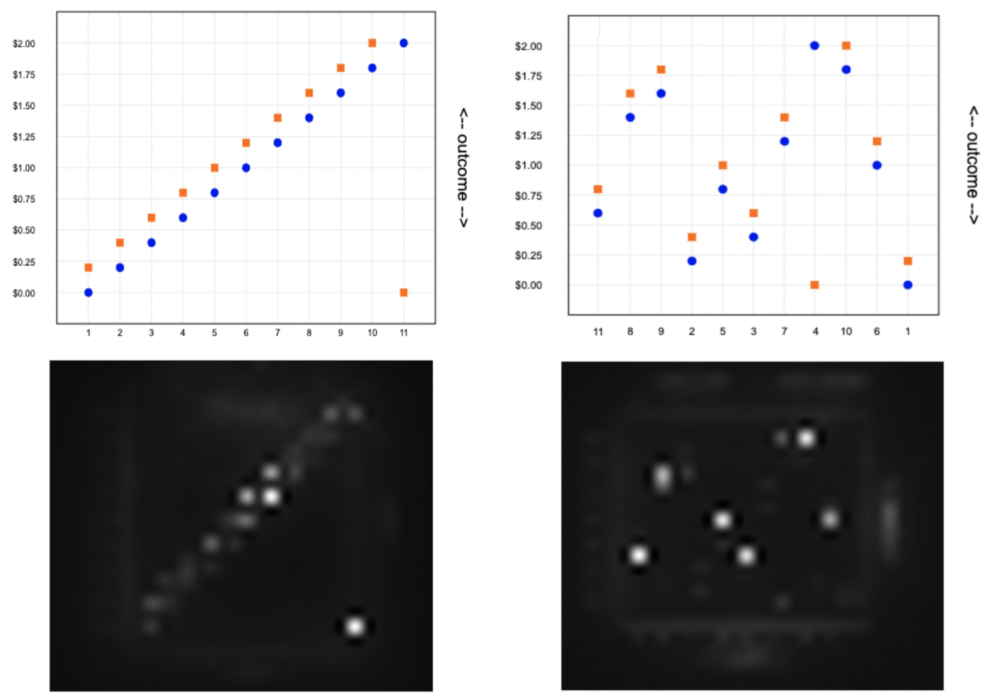
3.3.2 Manipulating Memory

Our first two memory manipulations sought to decrease the impact of imperfect recall on choice. The No Recall Bias treatment proceeded in the same way as the Baseline Sequential Information treatment, except that participants ($N = 100$) were asked to write down outcomes during the sequential sampling process. Specifically, they were asked to have a pen and paper (or other method of recording) ready and incentivized with a potential bonus payment for correctly responding to a quiz at the end of the round (see Figure IA3). By reducing the scope for imperfect recall,

[19] We still display the verbal description of all outcome pairs and probabilities, as in the Baseline Simultaneous Information treatment.

[20] In short, the model generates a saliency map of an image using an algorithm that was validated through measures of human visual fixation. The output of the saliency map is based on levels of activation in units of pixels—higher pixels/brighter areas are associated with higher levels of salience.

Figure 4: This figure depicts the relative salience of the unlikely-outperforming state of Asset U in the Baseline simultaneous information (left) versus No Salience Bias simultaneous information (right) treatments. The top two graphs correspond to the distributions shown to participants. The bottom two figures correspond to the salience maps of those distributions as measured by the Graph-Based Visual Saliency Method. In the Baseline simultaneous information treatment, the performance of Asset F in state 11 was by far the most salient area, with an average activation level of approximately 230 pixels (the next highest outcome had an activation level of less than 100). In contrast, in the No Salience Bias simultaneous information treatment, the outcomes associated with state 11 had an average activation level of approximately 30 pixels, while six other states had outcomes in the range of 100 to 150 pixels.



participants in the No Recall Bias treatment are predicted to display a smaller preference for Asset U than those in the Baseline Sequential Information treatment.

Based on the comparative static in Prediction 2, the Minimal Requirement treatment reduced the number of states to potentially remember from 11 to 3—a reduction below critical thresholds for memory capacity (e.g., Miller, 1956). This represents the minimal example where it is still possible to have "frequent outperformance" of Asset F (in 2 of 3 states) while maintaining the same marginal distributions for both assets ($0, $1, and $2 with equal likelihood). Reducing the number of states decreases the likelihood that 1) the state where Asset U outperforms is misremembered while 2) a state where Asset F outperforms is correctly remembered. In turn, participants ($N = 99$) in the Minimal Requirement condition are predicted to prefer Asset U to a smaller extent than those in the Baseline Sequential Information treatment.

Our third manipulation sought to test Prediction 4 directly. Here, we exploited the role of interference in an attempt to manipulate memory formation in the Sequential Information learning environment. To that end, we introduced a decoy Asset D which could never be chosen; the choice remained the same—between Asset U and F. Participants sequentially sampled outcomes from Asset D along with the two other assets.[21] The decoy asset's outcomes dominate the outcomes of the two assets that can be chosen, so that it should be perceived as very attractive. Prediction 4 says that assets which are similar to the decoy should also be perceived as more attractive due to interference in memory formation, even if the features that create similarity are payoff-irrelevant.

Using color (shades of orange vs. blue) and industry classification (oil vs. tech), we structured the decoy Asset D to either be similar to Asset U or F. Figure 5 shows the conditions where Asset D was similar to Asset U ("Interference (Unlikely-Outperformance)" condition) and where Asset D was similar to Asset F ("Interference (Frequent-Outperformance)" condition). Participants were randomized into the former ($N = 99$) or the latter condition ($N = 102$). While participants in the Interference (Frequent-Outperformance) condition are predicted to continue selecting Asset F over U, the framework predicts that not only will the gap decrease in the Interference (Unlikely-Outperformance) condition, but it may actually reverse.

4 Empirical Results

This section reports the results on choices (Section 4.1) and beliefs (Section 4.2). We explore additional mechanisms in Section 4.3. We start with our Baseline Experiment and then proceed with the theory-testing conditions.[22]

4.1 Choices

Figure 2 presents results across the Baseline and theory-testing conditions. Table 1 shows our results on choices across the learning environments, with the Baseline conditions marked by E1

[21] See Table IA2 for the full joint distribution.

[22] We present the 11-state version of the experiment here. As there were no significant differences in the order in which beliefs and choices were elicited, we pool across these variations when presenting our main results. The Internet Appendix contains additional results and sample splits.

Figure 5: This figure illustrates the two interference conditions. Participants sequentially sampled an additional Asset D but could only select between the other two assets, which exhibit the same marginal distributions as Assets U and F in the Baseline design. Asset D dominates the two assets that can be chosen. In the 'Unlikely-Outperformance' variation (a), the dominant Asset D is similar (in color and industry) to the unlikely-outperforming Asset U ('Tech Company 2'). In the 'Frequent-Outperformance' variation (b), Asset D is similar to the frequently-better Asset F instead ('Oil Company 2'). Note that the ordering of assets was randomized.

(a) Interference (Unlikely-Outperformance) **(b)** Interference (Frequent-Outperformance)

and D1. Notably, only 5% of choices corresponded to indifference; removing those choices from the analysis does not change the results neither qualitatively nor statistically.

Table 1: Main Results – Choices. In this table, we report average choices. We report the average fraction of participants selecting each of the alternatives. t-statistics are based on participant-clustered standard errors whenever both rounds are included. */**/*** indicates statistical significance at the 10%/5%/1% level.

Condition	Choice Asset U	Choice Asset F	Difference versus 50%	t-test	N
Simultaneous Information					
(D1) Baseline	58.01%	41.99%	-8.01%***	-3.16	381
(D2) No Salience Bias	34.00%	66.00%	16.00%***	3.36	100
Sequential Information					
(E1) Baseline	15.10%	84.90%	34.90%***	19.57	404
(E2) No Recall Bias	37.00%	63.00%	13.00%***	2.68	100
(E3) Minimal Requirement	37.37%	62.63%	12.63%***	2.58	99
(E4.1) Interference (Frequently-Outperforming)	19.61%	80.39%	30.39%***	7.69	102
(E4.2) Interference (Unlikely-Outperforming)	67.68%	32.32%	-17.68%***	-3.74	99
Difference Across Learning Environments					
(E1-D1) Baseline			42.91%***	13.98	785
(E2-D2) Eliminating the Gap			-3.00%	-0.44	200

Consistent with our first Prediction 1, people learning about the distribution in the Simultaneous Information condition (D1) had a significant preference for the *unlikely-outperforming* Asset U, choosing it 58.01% of the time ($p < .01$).[23] On the other hand, consistent with Prediction 2, learning the same information through sequential sampling (E1) shifted choices towards the *frequently-outperforming* Asset F, which was selected 84.90% of the time ($p < .01$).

The gap in choices between the Simultaneous and Sequential Information environments (E1-D1) is substantial, at 42.91 percentage points ($p < .01$). The existence of the gap is surprising given that the underlying distributions and information about the assets were kept constant, which rules out many potential mechanisms such as prospect theory or sampling error.

We next vary the Baseline Experiment's design to exogenously manipulate memory and attention. Decreasing the salience of the state where Asset U outperforms (condition D2) flips the preference for the unlikely-outperforming asset towards the frequently-outperforming: the fraction of participants selecting Asset F increases from 41.99% to 63.00% ($p < .01$). The elimination—and even a reversal—of the preference for the asset with the unlikely upside suggests that attention and salience play an integral role in driving decisions from Simultaneous Information.

Similar to decreasing salience bias with treatment D2, we sought to reduce the bias induced by imperfect recall in the sequential learning environment with the No Recall Bias treatment. Row (E2) of Table 1 shows that decreasing the scope for imperfect recall reduced the fraction of choices for the frequently-outperforming Asset F from 84.90% to 63.00% ($p < .01$ both between treatments and relative to 50%). Notably, comparing the No Salience Bias and No Recall Bias treatments (E2-D2) shows the complete elimination of the choice gap between learning environments: people's choices were essentially the same regardless of whether they learned from simultaneous (D2) or sequential information (E2). This highlights the role of attention and memory constraints as the drivers of the gap in the Baseline conditions, rather than something inherent to the learning environment per se.

Next, as a separate test of the imperfect recall channel, we reduced the number of states from 11 to 3 in the Minimal Requirement treatment. According to Prediction 2, the lower number of states that need to be recalled should reduce the memory bias, and in turn, the preference for the frequently outperforming Asset F. Row (S3) of Table 1 shows that this is indeed the case: Asset

[23] Unless otherwise noted, p-values correspond to the difference from 50%.

F was chosen more than 20% less in the Minimal Requirement treatment than in the Baseline treatment (62.63%; $p < .01$ both between treatments and relative to 50%).

Finally, we move to testing the memory mechanism by manipulating interference in the sequential learning environment (Prediction 4). When the decoy asset is more similar to the frequently-outperforming asset in the Interference (Frequent-Outperformance) treatment, this merely reinforces the preference for Asset F. Indeed, as shown in row (E4.1) of Table 1, people select this asset at around the same rate as the the Baseline Sequential Information condition (E1). However, this preference changes dramatically when the decoy asset is more similar to the unlikely-outperforming asset. As shown in row (E4.2), the preference for Asset U goes from 15.10% in the Baseline Sequential Information condition to 62.89% in the Interference (Unlikely-Outperformance) condition.[24]

Adding a decoy asset that differed on non-payoff relevant attributes shifted choices from the vast majority of people preferring the frequently-outperforming Asset F to most people selecting the unlikely-outpeforming Asset U—a preference reversal of more than 40 percentage points! Manipulating similarity and interference led people to prefer the asset with *unlikely* outperformance in the sequential learning environment.

4.2 Beliefs

We now turn to examining beliefs. In line with the conjecture that the learning environment affects choices by changing people's perceptions of uncertainty, we find that beliefs largely follow the patterns in Section 4.1. Table 2 shows our main results on beliefs after learning in the simultaneous versus sequential learning environments. Beliefs about average outcomes are in Panel A and beliefs about the probability of outcomes above versus below \$1 are in Panel B. Overall, we find that participants' beliefs are consistent with their choices. Note that because the order of the belief and choice questions were randomized, these results are not driven by spillover effects, e.g., a desire to report beliefs that are consistent with one's choices.

In our Baseline condition, participants become highly over-optimistic about the performance of the frequently-outperforming asset after learning through sequential information (E1). The distribution they sampled was the same for both assets, with an average (and median) of exactly \$1. How-

[24] A simple dislike for a specific color or industry cannot explain this result, as it would suggest an equal reduction of the fraction of choices for Asset F in both (E4.1) and (E4.2).

Table 2: Main Results – Beliefs. In this table, we report average beliefs. Panel A reports the average outcome expected for each of the alternatives (both marginal distributions have an expected value of $1.00). Panel B reports the average probability expected for below/above-median outcomes for each of the alternatives (both marginal distributions have a 5-in-11 chance for these outcomes). The t-statistics in Panel B are from a difference-in-difference estimate between the below- and above-median probability assessments. t-statistics are based on participant-clustered standard errors whenever both rounds are included. */**/*** indicates statistical significance at the 10%/5%/1% level.

Condition	Beliefs Asset U	Beliefs Asset F	Difference (-in-Diff.)	t-test	N
Panel A: Expected outcome (0-2)					
Simultaneous Information					
(D1) Baseline	$1.07	$1.00	$-0.07***	-2.83	381
(D2) No Salience Bias	$1.06	$1.09	$0.03	0.81	100
Sequential Information					
(E1) Baseline	$1.00	$1.17	$0.16***	7.34	404
(E2) No Recall Bias	$1.08	$1.08	$0.00	0.06	100
(E3) Minimal Requirement	$1.09	$1.25	$0.16***	3.20	99
(E4.1) Interference (Frequently-Outperforming)	$0.99	$1.38	$0.38***	7.77	102
(E4.2) Interference (Unlikely-Outperforming)	$1.20	$1.08	$-0.12**	-1.97	99
Difference Across Learning Environments					
(E1-D1) Baseline			$0.23***	7.11	785
(E2-D2) Eliminating the Gap			$-0.03	-0.66	200
Panel B: Risk assessment (below/above-median probabilities)					
Simultaneous Information					
(D1) Baseline – P($x < 1$)	39.09%	44.27%	6.46%***	3.46	381
– P($x > 1$)	42.63%	41.35%			
(D2) No Salience Bias – P($x < 1$)	41.71%	42.39%	3.26%	1.14	100
– P($x > 1$)	44.16%	41.58%			
Sequential Information					
(E1) Baseline – P($x < 1$)	48.40%	38.24%	-25.87%***	-12.58	404
– P($x > 1$)	35.04%	50.76%			
(E2) No Recall Bias– P($x < 1$)	38.08%	40.55%	-0.68%	-0.20	100
– P($x > 1$)	37.87%	41.02%			
(E3) Minimal Requirement – P($x < 1$)	35.87%	31.82%	-9.39%**	-2.32	99
– P($x > 1$)	32.08%	37.42%			
(E4.1) Interference (Frequently-Outperforming) – P($x < 1$)	47.98%	37.38%	-39.47%***	-8.05	102
– P($x > 1$)	28.83%	57.70%			
(E4.2) Interference (Unlikely-Outperforming) – P($x < 1$)	34.88%	49.67%	13.66%**	2.54	99
– P($x > 1$)	42.35%	43.48%			
Difference Across Learning Environments					
(E1-D1) Baseline			-32.34%***	-11.60	785
(E2-D2) Eliminating the Gap			-3.94%	-0.87	200

ever, participants believed the average outcome is around \$1.17 for the frequently-outperforming Asset F, which is more than \$0.16 higher than their stated belief for the unlikely-outperforming Asset U ($p < .01$).[25] Consistently, they believe outcomes below \$1 are less likely (likelier) than outcomes above \$1 for the frequently-outperforming (unlikely-outperforming) asset, although all of these are the same (5 out of 11 situations below, 5 above \$1). The difference-in-differences measuring the overoptimism for the frequently-outperforming Asset F versus the unlikely-outperforming Asset U is 25.87 percentage points ($p < .01$). On the other hand, participants were more optimistic about the performance of the unlikely-outperforming Asset U in the Simultaneous Information treatment (D1). They expected the Asset U to return an average of \$0.07 more than Asset F ($p < .01$) and exhibit surplus probability mass of 6.46 percentage points for better outcomes of the unlikely-outperforming asset ($p < .01$).

There is a highly significant beliefs gap across the learning environments in our Baseline Experiment (E1-D1). Table 2's Panel A shows that this gap is \$0.23 for estimates of average outcomes ($p < .01$), i.e., participants expected an average outperformance of the frequently-outperforming asset over the unlikely-outperforming asset that is \$0.23 higher after learning from sequential information than from simultaneous information. Similarly, Panel B documents an excess probability mass of 32.34 percentage points ($p < .01$) for good versus bad outcomes of the frequently-outperforming asset relative to the unlikely-outperforming asset. To sum, the results on beliefs suggest that the systematic preference reversal in choices from Table 1 appears to be driven by a systematic change in perceptions—or mental representations—of uncertainty.

When we manipulate memory and attention, the gap in beliefs closes and then flips. We illustrate this for beliefs about average outcomes in (D2), (E2), and (E4.2).

First, decreasing salience bias in the Simultaneous Information treatment using the No Salience Bias condition led to more optimistic beliefs about Asset F by \$0.03, though the difference was not significant. Similarly, decreasing recall bias in the Sequential Information treatment with the No Recall Bias condition led to no significant differences in beliefs across the two assets. Attenuating biases in attention and memory closed the gap in beliefs as well as choices.

In the Interference (Unlikely-Outperformance) treatment, the decoy is predicted to interfere with recall about the unlikely-outperforming Asset U. In turn, participants were overoptimistic about

[25] Unless otherwise noted, p-values correspond to the difference between assets.

Asset U (\$1.21) relative to Asset F (\$1.06) by \$0.14 ($p < .01$). This is a substantial reversal compared to the Baseline Sequential Information treatment where they were overoptimistic about Asset F compared to Asset U. The gap in beliefs flipped from \$0.23 ($p < .01$) in our Baseline conditions to \$-0.14 ($p < .05$) when we manipulate memory, even though the learning environment was kept constant.

4.3 Alternative Explanations

Despite presenting people with the same information in both the simultaneous and sequential learning environments, it may still be possible that people just did not attend to the state where Asset U outperformed in the latter. This would generate a gap in choices and beliefs through a mechanism similar to *sampling error* (Fox and Hadar, 2006) or *selective sampling* (de Palma et al., 2014). We analyze viewing times for each outcome pair as a measure of attention to explore this possibility in our Baseline Experiment.

First, we find that participants do pay significantly more attention to extreme outcome pairs—consistent with salience-driven attention. The average number of seconds participants take to view a moderate outcome pair is 1.87, whereas states with extreme differences receive 3.65 seconds of viewing time, or 1.78 seconds more attention on average ($p < .01$; , see Table IA3). This near-doubling of attention for extreme outcome pairs confirms that the associated states are salient and not overlooked when people learn from sequential information.

Second, we test whether increased viewing times for states with extreme outcome differences are reflected in choices and beliefs. As shown in Figure IA4, participants who look longer at extreme outcome pairs compared to the average moderate outcome pair do not exhibit a weaker or stronger tendency to choose the frequently-outperforming Asset F compared to participants who do not pay more attention to these states. Table IA4 reports a regression showing that attention-weighted sampled outcome levels do not significantly predict choices. In summary, it seems that while salience does drive attention to states with extreme differences in the sequential learning environment, the bias generated by imperfect recall—as outlined in Prediction 2—plays the dominant role in affecting choices and beliefs.

5 Discussion

We find that the way that people learn about uncertainty changes their perception of it. People who learn from simultaneous information tend to overweight unlikely but salient outcomes while those who learn the same information sequentially tend to underweight such outcomes. These differences in mental representations are reflected in choices: learning about outcomes from "simultaneous" information leads to a preference for unlikely-outperforming assets, while learning the same information sequentially leads to a preference for frequently-outperforming assets. In line with our proposed framework, we demonstrate that the observed differences in beliefs and choices are driven by memory and attention constraints. This shows that the previously-documented "description-exprience" gap in risky choice can be derived from first principles of bounded rationality; manipulating factors related to cognitive constraints can close the gap completely, or even reverse it.

Our findings highlight how the same exact information can lead to vastly different perceptions of uncertainty depending on how it is learned. The kinds of unlikely but salient states that characterize the distributions in our experiment are not unusual, but rather common characteristics in a variety of settings. Fat tails and joint outliers are ubiquitous in a strongly connected society. The outsized influence of such outliers on individual and societal outcomes makes a better understanding of under- versus over-weighting of extreme states even more important.

Learning from sequential personal experiences has already been linked to choices of consumers and investors (Malmendier and Wachter, 2021). Our findings link this experience literature to a broader framework, which allows for a richer set of predictions on how behavior changes as a function of the learning environment and cognitive constraints. Integrating learning from sequential experiences into a common framework is also important because while the majority of studies have used "simultaneous" information about uncertainty—and developed models to fit the observed behavior (e.g., Prospect Theory)—learning from sequential information is an ecologically common way of learning about uncertainty.

Incorporating bounded attention and memory is also important because it helps to explain heterogeneity in risky choice (e.g., Diecidue et al., 2020). In a given learning environment, different subsets of participants may (i) correctly recall unlikely states and overweight them due to their salience, (ii) mis-remember unlikely states and neglect them (as if they "maximize the probability to be ahead"), (iii) or neither overweight nor underweight them (if the cognitive constraints are not binding).

Lastly, our results point to ways in which the learning environment can be tailored to improve the accuracy of mental representations of uncertainty. In our experiments, both the sequential and simultaneous learning environments led to biased beliefs. Exploring presentation formats that lead to more accurate beliefs is a promising avenue for research. Experimental research in the investment context already shows encouraging results. Kaufmann et al. (2013) document that a sampling procedure can be used to debias beliefs about returns of a single risky asset. In a context with multiple risky assets, Laudenbach et al. (2022) find that sequentially sampling return pairs instead of directly receiving simultaneous information about the joint return distribution improves beliefs about dependence and makes portfolio choice more consistent with normative models (Markowitz, 1952). Our framework on how memory and attention affect choices could help structure the search for effective learning environments tailored to specific contexts (see the literature on nudging and boosting, e.g. Hertwig and Grüne-Yanoff, 2017).

References

Anderson, John R, Dan Bothell, Christian Lebiere, and Michael Matessa, 1998, An integrated theory of list memory, *Journal of Memory and Language* 38, 341–380.

Ba, Cuimin, J Aislinn Bohren, and Alex Imas, 2022, Over-and underreaction to information, *Available at SSRN* .

Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *The American Economic Review* 98, 2066–2100.

Barberis, Nicholas, Abhiroop Mukherjee, and Baolian Wang, 2016, Prospect theory and stock returns: An empirical test, in *Review of Financial Studies*, volume 29, 3068–3107 (Oxford University Press).

Barberis, Nicholas C, and Lawrence J Jin, 2023, Model-free and model-based learning as joint drivers of investor behavior, Technical report, National Bureau of Economic Research.

Bordalo, Pedro, John J Conlon, Nicola Gennaioli, Spencer Y Kwon, and Andrei Shleifer, 2022a, Memory and Probability, *The Quarterly Journal of Economics* .

Bordalo, Pedro, John J Conlon, Nicola Gennaioli, Spencer Yongwook Kwon, and Andrei Shleifer, 2021, Memory and Probability.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Salience theory of choice under risk, *Quarterly Journal of Economics* 127, 1243–1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2013, Salience and Asset Prices, *American Economic Review* 103, 623–628.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2020, Memory, attention, and choice, *The Quarterly journal of economics* 135, 1399–1442.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2022b, Salience, *Annual Review of Economics* 14, 521–544.

Bose, Devdeepta, Henning Cordes, Sven Nolte, Judith Christiane Schneider, and Colin Farrell Camerer, 2022, Decision weights for experimental asset prices based on visual salience, *The Review of Financial Studies* 35, 5094–5126.

Boyer, Brian H., and Keith Vorkink, 2014, Stock options as lotteries, *The Journal of Finance* 69, 1485–1527.

Brandstätter, Eduard, and Christof Körner, 2014, Attention in risky choice, *Acta psychologica* 152, 166–176.

Bruce, Neil DB, and John K Tsotsos, 2009, Saliency, attention, and visual search: An information theoretic approach, *Journal of vision* 9, 5–5.

Bryzgalova, Svetlana, Anna Pavlova, and Taisiya Sikorskaya, 2023, Retail trading in options and the rise of the big three wholesalers, *The Journal of Finance* 78, 3465–3514.

Célérier, Claire, Gordon Liao, and Boris Vallée, 2021, The Price Effects of Innovative Security Design.

Célérier, Claire, and Boris Vallée, 2017, Catering to investors through security design: Headline rate and complexity, *Quarterly Journal of Economics* 132, 1469–1508.

Chen, Daniel L, Martin Schonger, and Chris Wickens, 2016, oTree-An open-source platform for laboratory, online, and field experiments, *Journal of Behavioral and Experimental Finance* 9, 88–97.

Chesney, Marc, Felix Fattinger, and Jonathan Krakow, 2021, Interest Rates, Bounded Rationality, and Complexity: Demand and Supply of Retail Financial Products.

Chetty, Raj, Adam Looney, and Kory Kroft, 2009, Salience and taxation: Theory and evidence, *American economic review* 99, 1145–1177.

Cosemans, Mathijs, and Rik Frehen, 2021, Salience theory and stock prices: Empirical evidence, *Journal of Financial Economics* 140, 460–483.

Cubitt, Robin, Orestis Kopsacheilis, and Chris Starmer, 2022, An inquiry into the nature and causes of the description-experience gap, *Journal of Risk and Uncertainty* 65, 105–137.

d'Acremont, Mathieu, and Peter Bossaerts, 2016, Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response: Supplementary info 1, *Cerebral Cortex* 26, 1818–1830.

Danz, David, Lise Vesterlund, and Alistair J Wilson, 2022, Belief elicitation and behavioral incentive compatibility, *American Economic Review* 112, 2851–2883.

de Palma, André, Mohammed Abdellaoui, Giuseppe Attanasi, Moshe Ben-Akiva, Ido Erev, Helga Fehr-Duda, Dennis Fok, Craig R Fox, Ralph Hertwig, Nathalie Picard, et al., 2014, Beware of black swans: Taking stock of the description–experience gap in decision under uncertainty, *Marketing Letters* 25, 269–280.

Diecidue, Enrico, Haim Levy, and Moshe Levy, 2020, Probability dominance, *Review of Economics and Statistics* 102, 1006–1020.

Enke, Benjamin, and Thomas Graeber, 2023, Cognitive uncertainty, *The Quarterly Journal of Economics* 138, 2021–2067.

Enke, Benjamin, Thomas Graeber, and Ryan Oprea, 2023, Complexity and hyperbolic discounting .

Enke, Benjamin, Frederik Schwerter, and Florian Zimmermann, 2020, Associative memory and belief formation, Technical report, National Bureau of Economic Research.

Erev, Ido, and Greg Barron, 2005, On adaptation, maximization, and reinforcement learning among cognitive strategies., *Psychological review* 112, 912.

Erev, Ido, Eyal Ert, Alvin E Roth, Ernan Haruvy, Stefan M Herzog, Robin Hau, Ralph Hertwig, Terrence Stewart, Robert West, and Christian Lebiere, 2010, A choice prediction competition: Choices from experience and from description, *Journal of Behavioral Decision Making* 23, 15–47.

Erev, Ido, Ira Glozman, and Ralph Hertwig, 2008, What impacts the impact of rare events, *Journal of Risk and Uncertainty* 36, 153–177.

Fox, Craig R, and Liat Hadar, 2006, "Decisions from experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004), *Judgement and Decision Making* 1, 159–161.

Frydman, Cary, and Milica Mormann, 2016, The role of salience and attention in choice under risk: An experimental investigation, *University of Southern California, Marshall School of Business* .

Graeber, Thomas, Christopher Roth, and Florian Zimmermann, 2023, Stories, statistics, and memory .

Gupta, Neeraja, Luca Rigotti, and Alistair Wilson, 2021, The Experimenters' Dilemma: Inferential Preferences over Populations.

Handel, Benjamin, and Joshua Schwartzstein, 2018, Frictions or mental gaps: what's behind the information we (don't) use and when do we care?, *Journal of Economic Perspectives* 32, 155–178.

Harel, Jonathan, Christof Koch, and Pietro Perona, 2006, Graph-based visual saliency, *Advances in neural information processing systems* 19.

Hartzmark, Samuel M, Samuel D Hirshman, and Alex Imas, 2021, Ownership, learning, and beliefs, *The Quarterly journal of economics* 136, 1665–1717.

Henderson, Brian J., and Neil D. Pearson, 2011, The dark side of financial innovation: A case study of the pricing of a retail financial product, *Journal of Financial Economics* 100, 227–247.

Hertwig, Ralph, Greg Barron, Elke U Weber, and Ido Erev, 2004, Decisions From Experience and the Effect of Rare Events in Risky Choice, *Psychological Science* 15, 534–539.

Hertwig, Ralph, and Till Grüne-Yanoff, 2017, Nudging and boosting: Steering or empowering good decisions, *Perspectives on Psychological Science* 12, 973–986.

Hills, Thomas T, and Ralph Hertwig, 2010, Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions?, *Psychological science* 21, 1787–1792.

Kahana, Michael Jacob, 2012, *Foundations of human memory* (OUP USA).

Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–291.

Kaufmann, Christine, Martin Weber, and Emily Haisley, 2013, The Role of Experience Sampling and Graphical Displays on One's Investment Risk Appetite, *Management Science* 59, 323–340.

Kendall, Chad, and Ryan Oprea, 2024, On the complexity of forming mental models, *Quantitative Economics* 15, 175–211.

Khaw, Mel Win, Ziang Li, and Michael Woodford, 2021, Cognitive imprecision and small-stakes risk aversion, *The review of economic studies* 88, 1979–2013.

Kumar, Alok, 2009, Who gambles in the stock market?, *The Journal of Finance* 64, 1889–1933.

Laudenbach, Christine, Michael Ungeheuer, and Martin Weber, 2022, How to Alleviate Correlation Neglect in Investment Decisions, *Management Science* .

Li, Xiaomin, and Colin F Camerer, 2022, Predictable effects of visual salience in experimental decisions and games, *The Quarterly Journal of Economics* 137, 1849–1900.

Malmendier, Ulrike, and Stefan Nagel, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?, *The Quarterly Journal of Economics* 126, 373–416.

Malmendier, Ulrike, and Jessica A Wachter, 2021, Memory of past experiences and economic decisions, *Available at SSRN 4013583* .

Manski, Charles F, 2004, Measuring expectations, *Econometrica* 72, 1329–1376.

Markowitz, Harry M, 1952, Portfolio Selection, *Journal of Finance* 7, 77–91.

Mather, George, 2006, *Foundations of perception* (Taylor & Francis).

Miller, George A., 1956, The magical number seven, plus or minus two: Some limits on our capacity for processing information., *Psychological Review* 63, 81ÔÇô97.

Oprea, Ryan, 2022, Simplicity equivalents, *Working Paper* .

Paterson, Iain, and Andreas Diekmann, 1988, A Paradox in Decision Theory and Some Experimental Results: The Relative Nature of Decisions, *Theory and Decision* 25, 107–116.

Payzan-LeNestour, Elise, and Michael Woodford, 2022, Outlier blindness: A neurobiological foundation for neglect of financial risk, *Journal of Financial Economics* 143, 1316–1343.

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance* 19, 425–442.

Spitmaan, Mehran, Emily Chu, and Alireza Soltani, 2019, Salience-driven value construction for adaptive choice under risk, *Journal of Neuroscience* 39, 5195–5209.

Taubinsky, Dmitry, and Alex Rees-Jones, 2018, Attention variation and welfare: theory and evidence from a tax salience experiment, *The Review of Economic Studies* 85, 2462–2496.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.

Tversky, Amos, and Itamar Simonson, 1993, Context-dependent preferences, *Management science* 39, 1179–1189.

Ungeheuer, Michael, and Martin Weber, 2021, The Perception of Dependence, Investment Decisions, and Stock Prices, *Journal of Finance* 76, 797–844.

Ungeheuer, Michael, and Martin Weber, 2023, The Frequency of Outperformance and Investment Decisions.

Vokata, Petra, 2021, Engineering lemons, *Journal of Financial Economics* 142, 737–755.