Information Partitioning, Learning, and Beliefs[*]

PASCAL KIEREN[†] LUKAS MERTES[‡] MARTIN WEBER[§]

November 7, 2025

ABSTRACT

We experimentally study how information partitioning affects learning and beliefs. Holding the informational content constant, we show that processing small pieces of information at higher frequency (narrow brackets) causes beliefs to become overly sensitive to recent signals compared to processing larger pieces of information at lower frequency (broad brackets). As a result, partitioning information in narrow or broad brackets causally affects judgments. As mechanism, we provide direct evidence that partitioning information into narrower brackets shifts attention from the macro-level to the micro-level, which leads people to overweight recent signals when forming beliefs.

Keywords: biased beliefs, information bracketing, learning

JEL Classifications: D9, D1, G4

[*] We are grateful to Philip Bond, Peter Bossaerts, Sam Hartzmark, Camelia Kuhnen, Do Lee, Suk Lee, Vincenzo Pezone, Emanuel Renkl, Michael Thaler, Stefan Trautmann, Giovanni Trebbi, Agnieszka Tymula, and Georg von Weizsäcker for helpful comments and discussion. We also thank participants of the EEA, the Helsinki Finance Summit, the BFWG, the FIRS, the AFA, the SJDM, the DGF, the SAFE Household Finance Workshop, the MBEES, the Experimental Finance Conference, the Boulder Summer Conference on Consumer Financial Decision Making as well as seminar participants at the University of Mannheim. Pascal Kieren gratefully acknowledges financial support by the German Research Foundation (DFG) under the grant number 544579747.

[†] WHU – Otto Beisheim School of Management, `pascal.kieren@whu.edu`.

[‡] University of Mannheim, `lukas.mertes@uni-mannheim.de`.

[§] University of Mannheim, `weber@uni-mannheim.de`.

1 Introduction

When evaluating services, goods, or assets, individuals rely on information to assess their quality or determine whether certain desired characteristics are met. At the same time, both individuals and information providers decide how much information to consume or to present at a time. For instance, information can be processed in large batches at infrequent intervals or in smaller amounts at higher frequency. Consider an individual collecting several pieces of information. One approach is to evaluate the joint distribution of all individual signals and update beliefs once. Another possibility – which may be more natural if information is observed sequentially – is to consider each piece separately and update beliefs multiple times. According to Bayes' rule, such *information partitioning* (i.e., grouping individual information signals into sets) should not affect beliefs, since the law of total probability solely relies on the joint distribution of evidence. However, although partitioning does not affect a Bayesian learner, nearly every non-Bayesian updating model rests on some implicit assumption about how people process sets of information signals (Benjamin et al., 2016; Cripps, 2018).

To build some intuition, consider the model of diagnostic expectations (Bordalo et al., 2016; Bordalo et al., 2018; Bordalo et al., 2019), which captures overreaction in expectations driven by the representativeness heuristic (Kahneman and Tversky, 1972). In the model, subjective expectations are formed as a combination of rational expectations and a surprise component that overweights recent news. When forecasting future states, agents assign higher probability to states that are most representative based on recent news. Judgments of representativeness, however, may be influenced by how agents group or process sets of information signals, as smaller batches may convey a different impression than the aggregate.[1] If beliefs depend on how information is partitioned, multiple updated beliefs exist and additional assumptions are required to pin them down.

In this paper, we conduct a series of controlled experiments to examine whether information partitioning has a causal influence on learning and beliefs. While the influence of partitioning

[1] For instance, a price path with several short-term drops can still show an upward trend overall. While the overall upward-trending price path may be considered representative of a boom market, intermediate sequences of price drops may be more representative of a bust market.

received much attention for understanding how people make choices (Tversky and Kahneman, 1981; Read et al., 1999; Kőszegi and Matějka, 2020; Ellis and Freeman, 2024), its implications have not been studied for how people process information.[2] A major hurdle in studying the effect of information partitioning on beliefs is to find exogenous variation in the frequency and grouping of new information while holding the actual informational content constant. In controlled experiments, however, objective information and partitioning of information can be decoupled and exogenously varied. Additionally, experiments allow us to control subjects' information set, to specify the data-generating process, and to define subjects' payoff functions for reporting accurate expectations.

The experiments are organized around a stylized framework that builds on prior work by Gabaix (2019) and Enke and Graeber (2023). Subjects repeatedly observe signals – framed as price movements – to learn about the quality of a risky asset. Subjects know that the asset has a fixed probability of a price increase in each period, which we refer to as its fundamental quality. A price increase (decrease) corresponds to a positive (negative) signal about the asset's quality. Subjects' task is to infer the quality of the risky asset from observed price changes over 50 periods. A key feature of our experimental design is that we exogenously vary how information is partitioned. In the narrow information treatment, subjects observe each signal individually, which is implemented by observing a price chart slowly build over time. Every 10 periods, the build-up stops and subjects' beliefs about the asset's fundamental quality are elicited, reflecting information that is processed in smaller batches at higher frequency. In the broad information treatment, subjects receive all information at once, which is implemented by immediately observing the full price chart. Beliefs are elicited only after period 50, reflecting information that is processed in large batches at infrequent intervals. Importantly, once the price path in the narrow treatment is fully developed, the experimental interface looks identical across both treatments. As such, subjects have access to the same information and a Bayesian agent would provide identical posterior beliefs after the final period irrespective of how information is partitioned. This allows us to document systematic errors in the belief formation process which we can directly attribute to the partitioning of

[2] There is related evidence on myopic loss aversion showing that people are less risk-averse when outcomes are observed less frequently (Gneezy and Potters, 1997; Benartzi and Thaler, 1999). Our study differs in that we focus on beliefs rather than risk preferences.

information.

Our findings can be summarized as follows. We find that information partitioning significantly influences the belief formation process. Processing information in narrow brackets leads to greater estimation errors than processing information in broad brackets. The difference is sizable, as estimation errors in the narrow treatment are on average 28% greater than in the broad treatment. To understand why beliefs become less precise when information is processed in smaller partitions, we investigate the belief dynamics more closely. To do so, we follow Greenwood and Shleifer (2014) to estimate how information is weighted across treatments. We show that partitioning information in narrow brackets causes individuals to overweight recent information and to underweight distant information. These results are consistent with the well-established fact of belief extrapolation found in survey data and in asset prices (Barberis et al., 2015; Hirshleifer et al., 2015; Bordalo et al., 2019). This is not entirely surprising, since narrow bracketing is closely related to the standard paradigm of sequential belief updating. However, similar overweighting of recent information cannot be found when information is partitioned in broad brackets, where we find that individuals apply almost equal weight to all available information. As such, broad information partitioning appears to mitigate some well-documented judgmental biases.

Next, we aim to provide evidence on which psychological primitives drive the information partitioning effect. Drawing on recent research suggesting that attention and memory can generate belief extrapolation (Hartzmark et al., 2021; Enke et al., 2024; Jiang et al., 2024; Gödker et al., 2025), we conjecture that processing smaller pieces of information at higher frequency shifts attention from the macro-level (the joint informational content) to the micro-level (individual pieces of information). This heightened attention to small and frequent information signals causes beliefs to become overly sensitive to recent information which leads to extrapolation from such information. To test this mechanism, we proceed in two steps. First, we investigate subjects' memory across treatments. Attention and memory have a close relation as attention determines what information is encoded into memory (Kahana, 2012; Schwartzstein, 2014; Bordalo et al., 2020; Bohren et al., 2024). To analyze the influence of information partitioning on memory, we ask subjects to recall some of the encountered information. This task was unannounced beforehand and incentivized. Consistent with our conjecture, we find that subjects who process information at lower frequency have better memory of overall price-path characteristics. Additionally, we show that better memory also results in lower estimation errors.

Second, we aim to provide direct evidence for attention as underlying mechanism. We conduct another experiment (Experiment 2) in which we build on techniques from cognitive psychology to exogenously manipulate attention in the narrow information treatment (Verghese, 2001; Mrkva and Van Boven, 2017). Specifically, before reporting their final estimate in period 50, subjects have to watch the entire price path rebuild and have to identify the price of the asset for five randomly selected periods. Importantly, this manipulation does not provide any new information for subjects, but allows to examine whether an exogenous increase in attention to the overall information reduces the observed gap in beliefs. Our evidence is consistent with this conjecture, as the attention manipulation almost completely closes the gap in posterior beliefs between the narrow and broad treatment. Additionally, beliefs are not only less influenced by recent information, but subjects also recall the provided information better.

Finally, we explore the influence of information partitioning in two specific applications. First, we study an application to financial markets. A distinctive feature of financial markets is that not all available information (e.g., recent and past asset prices) is equally informative. If markets are efficient, stock prices will rapidly incorporate all value-relevant signals and thus quickly become stale information. This feature is different from the environment studied in our baseline experiments, in which all available information is equally important. To incorporate this, we conduct a third experiment (Experiment 3) in which price changes are governed by a Markov chain as in Frydman et al. (2014) and Charles et al. (2024). As a consequence, a Bayesian would overweight recent signals as they are more diagnostic of the underlying state. Consistent with our main result, we still find that subjects in the narrow treatment put more weight on recent information. However, by design, this behavior is now beneficial. Subjects in the narrow treatment now provide more accurate estimates than subjects in the broad treatment.

In a second application, we turn to a consumer choice setting. In online market places, consumers typically learn about the quality of a good by reading reviews from other customers. These reviews are often qualitative in nature, as customers write about their experiences and opinions. In a fourth experiment (Experiment 4), we test whether the partitioning of qualitative information can affect consumer judgments. In the experiment, subjects learn about the quality of a fictional smartphone by reading generic customer reviews and are asked to assess the phone's quality. Even in this setting, we find that narrow information partitioning leads to overweighting of recent information, while broad partitioning leads to equal weighting of information. Overall, these applications demonstrate that our results are highly consistent across information types (qualitative vs. quantitative), visual presentation (price chart vs. list), settings (financial markets vs. online market places), and data-generating processes (equal-weighted information signals vs. Markov chain).

Our findings add to the literature on behavioral biases in belief formation, as recently reviewed in Benjamin (2019). Prior research shows that people tend to neglect base-rates (Kahneman and Tversky, 1973; Fischhoff and Bar-Hillel, 1984), display overconfidence (Moore and Healy, 2008), do not sufficiently account for correlations in the data-generating process (Enke and Zimmermann, 2019; Ungeheuer and Weber, 2021), sometimes overinfer (Bordalo et al., 2018, 2020; Hartzmark et al., 2021; Kieren and Weber, 2025), and sometimes underinfer from recent signals (Edwards and Phillips, 1964; Phillips and Edwards, 1966). More recent research also investigates how heterogeneity in the learning environment can affect the belief formation process. Ba et al. (2024) and Augenblick et al. (2025) show that people overreact to information in complex and noisy environments, while they underreact in simple environments. Bohren et al. (2024) show that learning differs depending on whether information is acquired from descriptions or from sequential sampling. In a similar spirit, our main emphasis is not on the type of information being provided. Instead, we are interested in whether partitioning the same information into narrower or broader brackets affects judgment. An important conceptual question in sequential belief updating is how individuals group signals. If people treat individual signals as distinct samples, they would iteratively update their beliefs after each signal. Alternatively, if people pool all observed signals, they would update from their initial prior using the combined sample. As argued by Benjamin et al. (2016), differences in grouping can be a mechanism behind dynamically inconsistent behavior.[3] Despite its significance, this question has thus far received attention only in theoretical models (Benjamin et al., 2016; Cripps, 2018). Our results show that no specific grouping rule can be considered a universal feature of information processing. Instead, people group outcomes differently depending on how information is partitioned.

We also contribute to recent work studying the role of memory in belief formation (Mullainathan, 2002; Gennaioli and Shleifer, 2010; Malmendier et al., 2020; Wachter and Kahana, 2024; Bordalo et al., 2025). For example, Enke et al. (2020) show that people selectively recall pieces of information from the past if the context in which it is experienced is similar to today's context. Consistent with this notion, an increasing number of studies argues that selective recall of information might be a potential mechanism for self-servingly biased beliefs (e.g., Bénabou and Tirole, 2002; Chew et al., 2020; Zimmermann, 2020). Applied to investment decisions, Gödker et al. (2025) show that individuals tend to over-remember positive investment outcomes and under-remember negative ones. Jiang et al. (2025) find that investors are more likely to remember market episodes which are more similar to current market returns. Different from existing work, our study does not focus on the type of information that is being recalled (e.g., "good news" versus "bad news") but rather on how the partitioning of information makes the information more or less memorable. Information that is processed in different intervals – either due to contextual differences in information provision or due to individual preferences for consumption frequency – may thus be a source of persistent heterogeneity in expectations across individuals.

2 Conceptual Framework

In this section we develop a stylized framework to illustrate how information partitioning affects beliefs and to guide the design of our experiments. The mechanics build on prior work by Gabaix (2019) and Enke and Graeber (2023). The framework builds on two key features: (i) the belief formation process is not fully Bayesian; and (ii) information can be processed

[3] He and Xiao (2017) even show that assumptions on how people group signals will matter for *any* non-Bayesian updating rule.

either as one signal per period or in discrete batches of multiple signals every few periods.

Suppose in each period t, a good (or service) generates a binary information signal $s_t \in \{g, b\}$. Signals in all periods are i.i.d., and the full history of observed signals until period t is represented by $S_t = (s_1, s_2, \ldots, s_t)$. In financial markets, signals can be thought of as earnings surprises or dividends, whereas in consumer markets, signals can be thought of as quality signals. Which signal is generated in a given period depends on the good's underlying type. The agents' task is to observe the signals and to make inferences about the good's underlying type. The good can either be a good or a bad type, represented by G and B, respectively.[4] Each type corresponds to an underlying distribution from which signals are generated. The probability of observing signal g is θ_G for a good type, and θ_B for a bad type, with $0 < \theta_B < \theta_G < 1$. If signals are ordered outcomes (with good outcomes being preferred over bad outcomes), this implies that the good type has a signal distribution that first-order stochastically dominates the bad type. Let π_0^G and π_0^B represent the agent's prior belief about the good and bad type, respectively.

Following Gabaix (2019) and Enke and Graeber (2023), we assume that agents form beliefs based on a convex combination of their prior belief and the (Bayesian) posterior belief implied by the observed signal:

$$\pi_t^G(S_t) = (1-\lambda)\pi_{t-1}^G + \lambda \mu_t^G \tag{1}$$

with

$$\mu_t^G(S_t) = \frac{P(s_t \mid \theta_G)\pi_{t-1}^G}{\sum_{j=G,B} P(s_t \mid \theta_j)\pi_{t-1}^j}, \tag{2}$$

where $\pi_t^G(S_t)$ denotes the agent's posterior belief that the good's type is G conditional on the observed signal history S_t, π_{t-1}^G is the agent's prior belief in period $t-1$, μ_t^G is the posterior implied by the signal, $P(s_t \mid \theta_j)$ is the probability of observing signal s_t conditional on the type $j \in \{G, B\}$, and $0 \leq \lambda \leq 1$ represents the relative weight placed on the posterior implied by the observed signal. This model preserves the martingale property of the Bayesian

[4] While we focus on two types for tractability and illustration purposes, all implications hold for K types as well. The generic case of Equation 1 and Equation 2 is as follows: $\pi_t^i(S_t) = (1-\lambda)\pi_{t-1}^i + \lambda\mu_t^i$ and $\mu_t^i(S_t) = \frac{P(s_t|\theta_i)\pi_{t-1}^i}{\sum_{i\in K} P(s_t|\theta_i)\pi_{t-1}^i}$.

posteriors and nests Bayesian updating for $\lambda = 1$. It permits a variety of belief updates ranging from extreme dogmatism ($\lambda = 0$), where the agent entirely relies on their initial belief, to Bayesian updating ($\lambda \to 1$). The resulting updating process is comparable with an anchoring-and-adjustment heuristic (Tversky and Kahneman, 1974), in which individuals anchor their estimates on some initial belief and then adjust it in the direction of the observed information.

Next, assume that agents observe signals in batches of size $n \in \{1, 2, \dots\}$. Importantly, the good continues to generate a signal in each period t. However, agents may not observe the signal in each period. This could either be due to attentional constraints or due to deliberation. We denote the ordered set of signals that an agent observes in any period t by $s_{t:n} \equiv \{g_x, b_y\}^n$, where g_x and b_y denote the number of g and b signals contained in the batch, with $x + y = n$, and n denotes the size of the batch. Note that when an agent observes signal batch $s_{t:n}$ in period t, this implies that the last signal batch must have been observed in period $t - n$. As signals continue to be i.i.d., it follows that:

$$P(s_{t:n} \mid \theta_j) = \prod_{i=0}^{n-1} P(s_{t-i} \mid \theta_j). \tag{3}$$

We now consider how the agent's beliefs evolve over t periods under different assumptions regarding how information is incorporated into beliefs.

Narrow Information Processing. First, assume that information is observed and processed at higher frequencies. The limit case is $n = 1$. We call this *narrow information processing*. If information is processed in narrow batches of size 1, the agent iterates the updating process described in Equation (1) and (2) for each of the t periods:

$$\pi_t^{G,\text{narrow}}(S_t) = (1-\lambda)\pi_{t-1}^{G,\text{narrow}} + \lambda\mu_t^{G,\text{narrow}}, \tag{4}$$

with

$$\mu_t^{G,\text{narrow}}(S_t) = \frac{P(s_t \mid \theta_G)\pi_{t-1}^{G,\text{narrow}}}{\sum_{j=G,B} P(s_t \mid \theta_j)\pi_{t-1}^{j,\text{narrow}}}. \tag{5}$$

This updating procedure is the standard assumption in many models of rational belief updating, as new information is immediately incorporated into beliefs.

Broad Information Processing. Alternatively, assume that information is observed and processed at lower frequencies in batches of size $n > 1$. We call this *broad information processing.*[5] Since individuals cannot perform intermediate belief updates when observing information in batches, they perform a single belief update based on the prior they had in period $t - n$. Hence, the updating protocol is:

$$\pi_t^{G,\text{broad}}(S_t) = (1-\lambda)\pi_{t-n}^{G,\text{broad}} + \lambda \mu_t^{G,\text{broad}}, \tag{6}$$

with

$$\mu_t^{G,\text{broad}}(S_t) = \frac{P(s_{t:n} \mid \theta_G)\pi_{t-n}^{G,\text{broad}}}{\sum_{j=G,B} P(s_{t:n} \mid \theta_j)\pi_{t-n}^{j,\text{broad}}}, \tag{7}$$

where $P(s_{t:n} \mid \theta_j)$ is the probability of observing batch $s_{t:n}$ conditional on the type $j \in \{G, B\}$ as in Equation 3 and π_{t-n}^{G} is the agent's prior belief before observing the most recent batch in period $t - n$.

Note that the two updating processes outlined above impose different demands on attention and memory capacity. In both processes the current prior and the most recent signal batch are sufficient to know. However, depending on the process, the agent updates more (less) frequently and therefore must attend to a smaller (larger) batch size per update.

Importantly, the two processes also have different implications for agents' posterior beliefs, $\pi_t^{G,\text{narrow}} \neq \pi_t^{G,\text{broad}}$. The only exception is when $\lambda = 1$, i.e., belief updating is Bayesian (see Cripps, 2018; for a related discussion). Proposition 1 and 2 summarize this result. Proposition 1 states that information partitioning does not matter if the updating protocol is Bayesian. Proposition 2 summarizes the basic implication of information partitioning for non-Bayesian belief updating. It indicates that belief updating depends on how information is observed, as two different ways to process the same information until period t can generate different updated beliefs.

PROPOSITION 1: For $\lambda = 1$, partitioning information into smaller or larger brackets does not

[5] While we use $n_{narrow} = 1$ and $n_{broad} > 1$ for illustration purposes, it is sufficient to assume that $n_{narrow} < n_{broad}$.

influence posterior beliefs ($\pi_t^{j,\text{narrow}} = \pi_t^{j,\text{broad}}$). The proof of Proposition 1 is in Appendix A.1.

Proposition 2: For $\lambda \neq 1$, partitioning information into smaller or larger brackets generally results in different posterior beliefs ($\pi_t^{(j,\text{narrow})} \neq \pi_t^{(j,\text{broad})}$). The proof of Proposition 2 is in Appendix A.2.

Finally, Proposition 3 and 4 provide further comparative statics on the two updating protocols. First, note that although the two protocols will lead to different updated beliefs, there is no simple rule that describes which protocol leads to more optimism or pessimism. The reason is that narrow belief updating depends more heavily on the order of signals for $\lambda \neq 1$ than broad belief updating. Thus, depending on the signal ordering, beliefs in the narrow updating protocol can be both more optimistic or more pessimistic than in the broad updating protocol. Proposition 4 is a direct consequence. Since beliefs in the narrow updating protocol vary more across different signals orderings than in the broad updating protocol, the variance in belief is higher.

Proposition 3: For $\lambda \neq 1$, there is no simple comparative static that describes whether narrow or broad information processing makes individuals consistently more optimistic or pessimistic than the other. The proof of Proposition 3 is in Appendix A.3.

Proposition 4: For $\lambda \neq 1$, narrow information processing leads to a higher variance in posterior beliefs than broad information processing. The proof of Proposition 4 is in Appendix A.4.

Figure B1 in Appendix B illustrates the main results using simulated posterior beliefs according to narrow and broad information processing respectively. First, posterior beliefs in the narrow updating protocol are more dispersed than in the broad updating protocol. Second, posterior beliefs in the narrow updating protocol are sometimes above (more extreme) and sometimes below (less extreme) than in the broad updating protocol. This remains true even when controlling for the composition of signals.

3 Experimental Design

3.1 Baseline Design

In order to examine the causal effect of information partitioning on learning and beliefs we require a setting with the following features: (1) individuals repeatedly incorporate new information signals into their beliefs; (2) the frequency at which individuals process new information can be exogenously assigned; (3) beliefs can be compared to a normative benchmark; and (4) the belief elicitation is incentive-compatible. We design four preregistered experiments that accommodate these features.

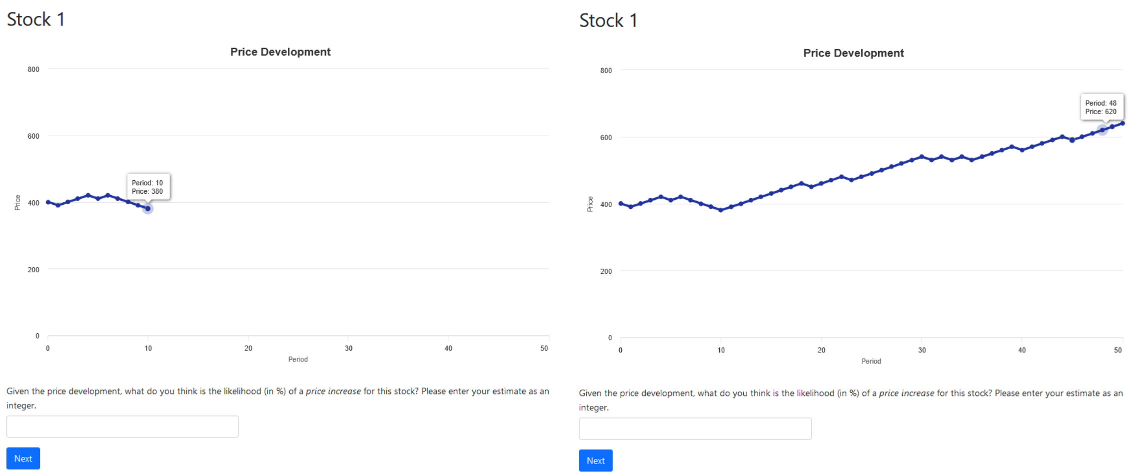
In this section, we outline the features of our baseline experiment (Experiment 1) in detail. In the experiment, subjects have to form beliefs about the fundamental quality of a risky asset. The asset has a fixed probability of a price increase, $s_i \in \{0.20, 0.21, ..., 0.80\}$, which represents its fundamental quality. The asset starts with an initial price of 400. In each period $t \in \{1, 2, ..., 50\}$, the price level of the asset either increases or decreases by a constant amount; a price increase is always 10 and a price decrease is always -10. In every period, the current and all prior price levels are provided to subjects in a price-line chart. Since a price increase is more likely to be observed if the risky asset has a higher fundamental quality s_i, price changes correspond to signals about the asset's fundamental quality.

Subjects are informed that the risky asset has a fixed fundamental quality that is uniformly distributed between 20% and 80%, but are not informed about the actual quality. Their task is to infer this quality from the observed price changes. Specifically, we ask subjects to report their belief p_i regarding the probability of a price increase of the risky asset. Subjects record their beliefs using a slider ranging from 20% to 80% in 1 percentage points increments. The key component of our study is whether the elicited beliefs regarding the fundamental quality depends on how information is partitioned. We introduce two between-subject treatments for eliciting beliefs at different frequency, termed *narrow information* and *broad information*.

In the *narrow information* treatment, we elicit subjects' beliefs about the asset's fundamental quality every 10 periods. This treatment aims to elicit beliefs when information is processed in small amounts at higher frequency. Subjects start with an empty price-line chart at period $t = 0$ which *builds* over time, reflecting the notion that small bits of information are observed at higher frequency. Once they are ready to begin with the task, they observe a price change every 0.5 seconds until they observe a total of 10 price changes, i.e., until period 10. Afterwards, we elicit their beliefs about the asset's fundamental quality (Figure 1a). This process then continues to build the price-line chart in batches of 10 periods[6] until subjects reach period 50, leading to a total of five estimates.

In the *broad information* treatment, we elicit subjects' beliefs about the asset's fundamental quality only once. This treatment aims to elicit beliefs when information is processed in large amounts at lower frequency. Similar to before, subjects start with an empty price-line chart at period $t = 0$. In contrast to the *narrow information* treatment, however, the graph does not build over time. Once subjects are ready to begin, they observe all price changes between period 1 and 50 at once. Afterwards, we elicit their beliefs about the asset's fundamental quality (Figure 1b).

Figure 1: Experiment 1: Treatments



(a) Narrow Information **(b)** Broad Information

Note: This figure presents exemplary screens of the estimation task as seen by subjects in Experiment 1. In the narrow information treatment (a), the price-line chart builds over time and subjects beliefs about the asset's fundamental are elicited every 10 periods. In the broad information treatment (b), subjects observe all price changes between period 1 and 50 at once and their beliefs are only elicited in period 50.

Overall, subjects play multiple trials each consisting of 50 periods. The fundamental

[6] Subjects always observe all past price changes from previous batches in the price-line chart.

quality of a risky asset is only fixed for one particular trial and as such varies across trials. This information is known to subjects. To account for the fact that subjects in the broad information treatment make fewer choices than those in the narrow information treatment, the former will complete eight trials while the latter only complete four trials. The experiment concludes with a brief survey about subjects' socioeconomic background.

To analyze the influence of information partitioning on beliefs, it is crucial for our experimental design that subjects across treatments have access to the same information. We address this in two ways. First, our main variable of interest is subjects' reported belief about the fundamental quality of the risky asset in period 50. This ensures that the available information set at the time of the decision is identical and allows to attribute any observed difference in beliefs to our treatments. Second, we follow convention in randomly generating the price paths before the experiment (e.g., Hartzmark et al., 2021; Fischbacher et al., 2017). This not only facilitates between-subject analyses but also allows a direct comparison of beliefs across treatments conditional on observing the same information. We first randomly drew 4 price paths for fundamental qualities greater than 50% ("positive paths" hereafter). Next, we mirrored each price path to obtain another 4 price paths for fundamental qualities of less than 50% ("negative paths" hereafter). This allows us to detect potential asymmetries between increasing and decreasing price paths, leading to 8 price paths. Finally, for each price path we rotated price changes to create variation in observed price path patterns without affecting the final price (and thus increasing statistical power without changing the final Bayesian posterior). This way, we generated a total of 24 price paths.

Subjects are incentivized based on the accuracy of their estimates. At the end of the experiment, we randomly select three estimates. For each selected estimate p_i that is within plus or minus 5 percentage points of the true probability of a price increase, s_i, subjects receive a bonus of £0.3. Additionally, subjects receive a fixed participation fee of £1.25. We chose this incentivization mechanism for its simplicity by imposing fewer cognitive burdens on subjects. Overall, this creates a simple and transparent learning environment that fosters truthful reporting as the number of price increases and decreases are a sufficient statistic for calculating the posterior probability. In contrast, recent studies show that more complex incentivization schemes such as the Binarized Scoring Rule (or variations such as the Quadratic Scoring Rule) can systematically bias truthful reporting, resulting in greater errors rates relative to simpler mechanisms (Danz et al., 2022).

In addition to the estimation task, we add a memory elicitation task in the spirit of Gödker et al. (2025) to control the influence of our treatments on subjects' memory. The memory elicitation consists of a number of questions in which subjects have to recall specific outcomes of the risky asset. Specifically, we ask subjects to recall how many positive and negative price changes they observed, the final price after period 50, as well as the maximum streak length of subsequent positive respectively negative price changes. The memory task always occurs after either the first or the last trial of an experiment in a counterbalanced order. The memory task is not announced beforehand and subjects no longer have access to the price-line chart. We aggregate the number of correctly recalled questions to an overall memory score ranging from 0 (none of the five questions was answered correctly) to 5 (all questions were answered correctly). The memory task is incentivized in addition to the estimates. Subjects receive £0.1 for each correctly recalled question.

3.2 Recruitment Procedure

The experiment was computerized using oTree (Chen et al., 2016). We recruited a total of 3,059 individuals ($N = 713$ for Experiment 1) from the crowdsourcing platform Prolific to participate in four experiments. The design, hypothesis, and sample selection criteria are all preregistered.[7] The study obtained ethics approval by the Institutional Review Board of the authors' institution. The subject pool is comprised of subjects from the UK and the US.

[7] The preregistration documents can be found at https://aspredicted.org/LJ1_GZ3 (Experiment 1), https://aspredicted.org/DJP_5ZZ (Experiment 2), https://aspredicted.org/rp22-jwdt.pdf (Experiment 3), and https://aspredicted.org/jhf2-wk5x.pdf (Experiment 4).

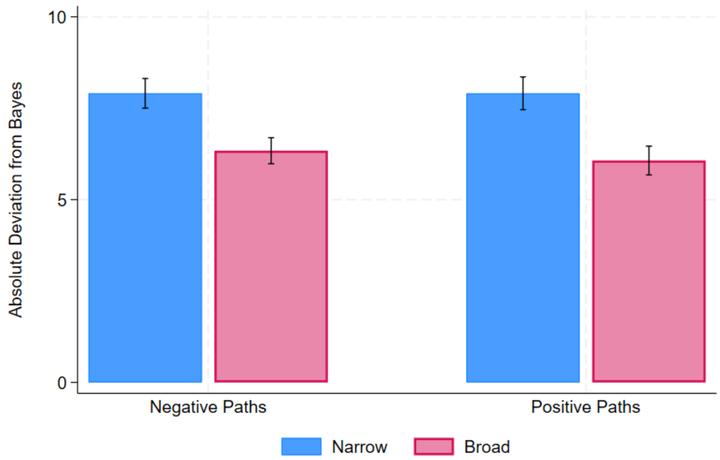
4 Results

4.1 Information Partitioning and Learning

Accuracy of Beliefs

To assess the influence of information partitioning on beliefs, we first compare subjects' estimation error, computed as the absolute difference between reported beliefs p_i and Bayesian beliefs b_i in period 50, between the narrow and broad information treatments. If information partitioning does not affect beliefs, there should be no difference in estimation error across treatments. Figure 2 plots the average estimation error for each treatment, split by positive and negative price paths.

Figure 2: Experiment 1: Estimation Error



Note: This figure plots the average estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) per treatment in Experiment 1 for negative respectively positive price paths.

First, note that irrespective of whether the price path was increasing or decreasing, the estimation error in the in narrow information treatment (blue bars) is significantly higher than the estimation error in the broad information treatment (red bars).[8] For negative price paths, the estimation error in the narrow treatment is 7.9 percentage points, while the estimation error in the broad treatment is 6.3 percentage points, leading to a difference of 1.6 percentage

[8] In fact, the estimation error in the narrow treatment is higher than the estimation error in the broad treatment for any Bayesian posterior in our experiment.

points ($p < 0.001$). For positive price paths, the estimation error in the narrow treatment is also 7.9 percentage points, while the estimation error in the broad treatment is 6.1 percentage points, leading to a difference of 1.8 percentage points ($p < 0.001$). In relative terms, this implies that observing information at higher frequency leads to estimation errors which are on average 25% and 30% larger relative to observing information at lower frequency.

While the pattern in Figure 2 provides first insights into the influence of information partitioning on beliefs, we test the following regression model to account for the dependence of observations:

$$|p_i - b_i| = \alpha + \beta_1 b_i + \beta_2 Narrow_i + \epsilon_i \tag{8}$$

We regress subjects' estimation error on the Bayesian posterior in period 50 and a narrow information dummy, which equals 1 if a subject is in the narrow information treatment and 0 otherwise.

Table 1: Experiment 1: Estimation Error

	Overall	Overall	Negative	Positive
bayes	-0.07	-0.09	0.25^{***}	-0.13^{***}
	(0.00)	(0.00)	(0.04)	(0.04)
narrow	1.70^{***}	1.67^{***}	1.51^{***}	1.82^{***}
	(0.29)	(0.28)	(0.33)	(0.37)
Controls	No	Yes	Yes	Yes
N	3,700	3,656	1,851	1,805
R^2	0.02	0.04	0.06	0.05

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, the narrow information dummy and control variables as independent variables for Experiment 1. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table 1 presents the results estimated on the whole sample (Columns (1) and (2)), as well as for positive and negative price paths separately (Columns (3) and (4), respectively). Across all specifications, the coefficient of $Narrow$ is positive and highly statistically significant ($p < 0.001$), confirming that information partitioning – consistent with Proposition 2 – has a causal effect on the formation of beliefs.

Information Weights and Belief Formation

To obtain a better understanding why subjects in the narrow information treatment report less accurate beliefs despite observing the same information, we investigate subjects' reaction to the provided information more closely. Specifically, we examine whether partitioning information into narrower or broader brackets affects how subjects weight information brackets when forming their beliefs. To do so, we follow Greenwood and Shleifer (2014) and regress subjects' final posterior belief p_i in round 50 on the weighted sum of past price changes.[9] In particular, we estimate nonlinear least squares regressions of the following form:

$$p_i = a + b \cdot \sum_{j=0}^{k} \omega_j \Delta Price_{t-j} + u_t, \tag{9}$$

with

$$\omega_j = \frac{(\delta_{broad} + \delta_{narrow} \cdot Narrow)^j}{\sum_{l=0}^{k} (\delta_{broad} + \delta_{narrow} \cdot Narrow)^l}.$$

We augment the model by Greenwood and Shleifer (2014) by adding an additional δ coefficient interacted with the narrow information dummy. The δ coefficients measures how much weight is placed on recent versus past brackets of information. δ_{broad} gives the baseline estimate for subjects in the broad treatment. δ_{narrow} estimates the difference between the narrow and the broad treatment. The sum of δ_{broad} and δ_{narrow} thus provides the estimate for subjects in the narrow treatment. The lower the estimate, the less important is past information relative to recent information for the posterior belief.

Column (1) of Table 2 presents the results. In Columns (2) and (3) we estimate the regressions without the treatment interaction separately for subjects in the narrow and broad information treatment. As can be inferred, subjects in the broad treatment place almost equal weight on all information brackets as a δ of 1 indicates no decline in weights. Notice that our design does not favor more recent information. Instead, a Bayesian agent would put equal weight on all encountered signals. Subjects in the broad treatment thus behave almost Bayesian. However, subjects in the narrow treatment exhibit a much smaller δ coefficient of 0.74. They put less weight on distant information and more weight on recent

[9] Note that in our setting price changes correspond to returns used by Greenwood and Shleifer (2014).

Table 2: Experiment 1: Information Weights

	(1)	(2)	(3)
δ_{broad}	0.97***	0.97***	
	(0.02)	(0.02)	
δ_{narrow}	-0.23***		0.74***
	(0.03)		(0.02)
a	48.60***	48.78***	48.44***
	(0.15)	(0.20)	(0.23)
b	0.41***	0.43***	0.41***
	(0.00)	(0.00)	(0.00)
N	3,700	1,696	2,004
R^2	0.82	0.86	0.79

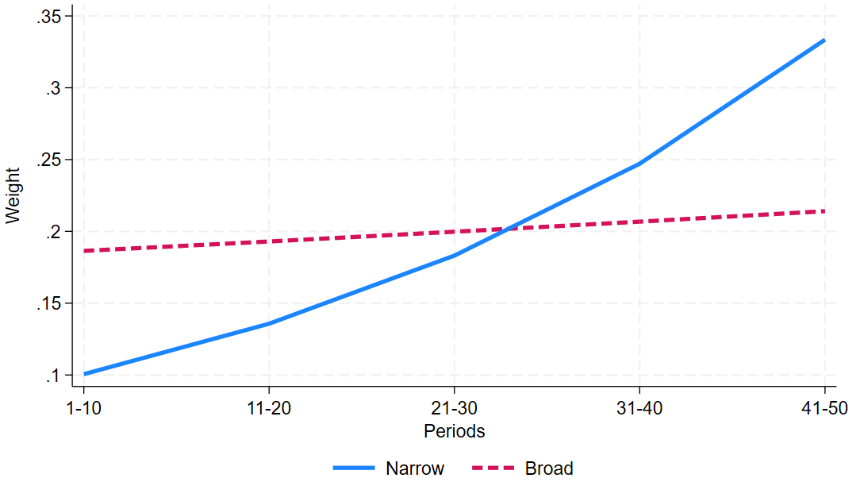
Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

information than subjects in the broad treatment. This difference is highly significant and economically meaningful as shown by the negative interaction ($p < 0.001$) in column (1). In other words, partitioning information into narrower brackets causes individuals to overweight recent information and to underweight distant information when forming beliefs.[10] Figure 3 illustrates the finding using the estimated δ coefficients of column (1). Applying the same methodology to simulated beliefs according to our framework results in an identical pattern (see B2 in Appendix B).

Since subjects in the broad treatment assign more equal weight to all available information, we should observe that their beliefs are also less influenced by the order of signals than the beliefs by subjects in the narrow treatment (compare Proposition 4). We test this by exploiting our design feature of rotating price changes without affecting the final price, resulting in 8 price path groups that each comprise 3 price paths with identical final price but different order of price changes. Figure 4 plots the standard deviation of estimates per treatment across all 8 price path groups. In each group, the standard deviation of beliefs is indeed larger in the narrow treatment than in the broad treatment.

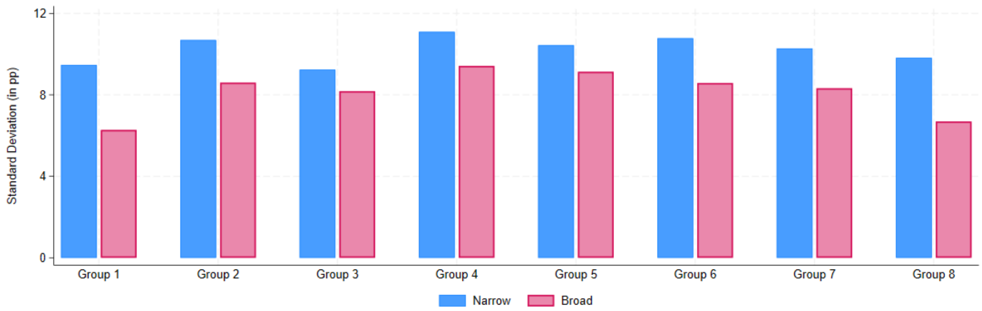
[10] In Table B1 in Appendix B, we also report alternative specifications based on the frequency of price increases instead of Greenwood and Shleifer (2014) regressions. All conclusions remain the same.

Figure 3: Experiment 1: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 9 in Experiment 1.

Figure 4: Experiment 1: Variation of Beliefs



Note: This figure plots the standard deviation (in percentage points) of posterior beliefs in period 50 across price path groups per treatment in Experiment 1.

4.2 Exploring the Mechanism

The previous section demonstrates that partitioning information into narrower or broader brackets influences learning and beliefs. In this section, we aim to provide evidence for a specific mechanism behind this effect. Note that the belief partitioning effect is not consistent with Bayesian learning which predicts no difference depending on how information is partitioned (see Proposition 1). Additionally, the effect is not consistent with models of motivated beliefs (e.g., Kunda, 1990; Brunnermeier and Parker, 2005) or misattribution (e.g., Ross, 1977; Gagnon-Bartsch and Bushong, 2022) that predict differences based on how desirable information is but not depending on how the same information is partitioned, or models of recency bias (e.g., Camerer and Hua Ho, 1999; Fudenberg et al., 2014) that predict *no* difference as the most recent information is identical across treatments.

As discussed in Section 2, our framework imposes different demands on attention depending on whether information is processed in narrow or broad brackets. Subjects in our broad treatment must attend to the informational content of the entire price path when forming their belief in period 50. For subjects in the narrow treatment, however, it is sufficient to attend to the most recent 10 period information bracket when providing their beliefs. We therefore consider a mechanism under which partitioning information into narrower brackets shifts attention from the macro-level (i.e., the joint informational content) to the micro-level (i.e., individual pieces of information). This heightened attention to small and frequent information signals causes beliefs to become overly sensitive, which leads to the overweighting of individual information brackets in the narrow treatment. The conjecture that increased attention to individual information fosters greater extrapolation is supported by recent studies in economics and cognitive psychology. For instance, Hartzmark et al. (2021) show that ownership channels attention towards signals associated with owned goods which leads to overextrapolation from such signals. Additionally, over-extrapolation is at least partly driven by the associative nature of memory through recall when making judgments (e.g., Gennaioli and Shleifer, 2010; Bordalo et al., 2020; Enke et al., 2020). For example, Enke et al. (2020) show that people overreact to information because they are more likely to recall similar prior information.

Prior work has shown that attention and memory have an intimate relation. Before a signal can be recalled, it must first be encoded into memory. In fact, attention determines what type of information is encoded into memory (Schwartzstein, 2014; Hartzmark et al., 2021; Bohren et al., 2024). If partitioning information into narrower or broader brackets affects attention and thus determines which signals are encoded into memory, then the cognitive process outlined above generates testable hypotheses between the partitioning of information and individuals' beliefs and memory. We conjecture that – in the spirit of choice bracketing (e.g. Read et al., 1999) – processing information at lower frequency (narrow information treatment) causes individuals to selectively focus their attention on small blocks of information, thereby losing sight of the big picture. As a result, beliefs become overly sensitive to recent information. We therefore expect that excessive attention to the micro-level leads to worse memory at the macro-level, which eventually causes beliefs to be further away from the Bayesian benchmark.

Information Partitioning and Memory

To test the mechanism outlined above, we first investigate whether partitioning information into narrower or broader brackets influences how memorable the observed information is. If subjects in the broad information treatment pay more attention to information at the macro-level (i.e., they pay equal attention to all information brackets) than subjects in the narrow information treatment, they should answer more memory questions correctly.

Table 3 displays subjects' answers to the memory questions elicited in the baseline experiment. Panel A shows the fraction of subjects who answered the respective question correctly. Across questions and treatments approximately 17% of questions were answered correctly. When the questions ask for the number of increases, the number of decreases, or the final price, the fraction of subjects who answered correctly is significantly higher in the broad information treatment than in the narrow information treatment. The differences are sizable as the share of correct answers in the broad information treatment is roughly 50% higher than in the narrow information treatment. There is no difference across treatments for the questions about the maximum streak length of increases and decreases. Comparing the memory score, i.e., the number of correctly answered questions, across treatments reveals a similar pattern (Panel B).

Our memory analysis shows that subjects in the broad information treatment exhibit a better recall than those in the narrow information treatment. Next, we investigate the transmission of memory on the formation of beliefs by regressing subjects' estimation error on their memory score. Table 4 displays results for the whole sample (Column 1), the whole sample with a treatment interaction (Column 2) and for each treatment separately (Columns 3 and 4). All specifications consistently show that the number of correctly answered

Table 3: Experiment 1: Memory

Panel A: Fraction	in %		
	Broad	Narrow	Difference
increases	19.81	13.57	6.24^{**}
			(2.11)
decreases	19.81	13.77	6.04^{**}
			(2.03)
final price	25.00	16.37	8.63^{***}
			(2.69)
streak up	16.04	17.96	-1.93
			(0.62)
streak down	17.45	17.17	0.29
			(-0.09)
Panel B: Number			
	Broad	Narrow	Difference
memory score (all 5)	0.98	0.79	0.19^{*}
			(1.83)
memory score (first 3)	0.65	0.44	0.21^{***}
			(3.53)

Note: This table displays answers to the memory questions of Experiment 1. Panel A displays the fraction of subjects who answered correctly per question and by treatment (broad vs. narrow information). Panel B displays the memory score out of all 5 and out of the first 3 questions by treatment. In both panels, the final column presents Mann-Whitney tests for differences in means across treatments; the corresponding z-scores are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

memory questions is negatively related to the absolute difference between subjective and objective beliefs. In other words, better recall at the macro-level leads to more accurate beliefs, consistent with the outlined mechanism. Importantly, the relation is present in both treatments and approximately equal in magnitude, suggesting that although memory differs across treatments, the transmission mechanism of memory on beliefs is not affected by how information is partitioned.

Information Partitioning and Attention

Finally, we aim to provide direct evidence for attention as underlying mechanism behind the belief partitioning effect. We use a comparative static approach to exogenously manipulate attention in the narrow information treatment. If the effects in our baseline experiment are driven by differences in attention depending on how information is partitioned, then

Table 4: Experiment 1: Memory and Accuracy

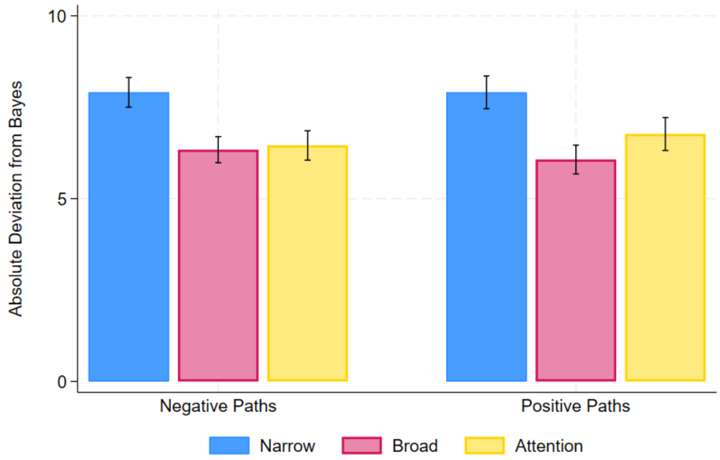
	Overall	Overall	Narrow	Broad
memory score	-0.92***	-0.93***	-0.91***	-0.93***
	(0.11)	(0.15)	(0.17)	(0.16)
narrow	1.53***	1.59***		
	(0.28)	(0.34)		
memory score × *narrow*		0.03		
		(0.23)		
N	3,700	3,700	2,004	1,672
R^2	0.05	0.05	0.02	0.04

Note: This table shows regressions with the absolute deviation of subjects' beliefs from the Bayesian posterior in period 50 as dependent variable and the memory score, and control variables as independent variables. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

exogenously manipulating attention in the narrow information treatment should reduce the observed gap in beliefs across treatments. To do this, we conduct an additional experiment (Experiment 2; $N = 352$), which directly builds on our baseline design. As in the narrow information treatment of the baseline experiment, subjects observe a price change every 0.5 seconds and report their beliefs about the asset's fundamental quality every 10 rounds. However, before reporting their final belief in period 50, subjects have to watch the entire price path rebuild and have to identify the price of the asset for five random periods (See Figure C10 in Appendix C.). This method is inspired by prior research in cognitive psychology, which shows that visual search fosters attention (Verghese, 2001; Mrkva and Van Boven, 2017). The periods are randomly drawn such that each 10-period bracket of the price path, i.e., periods 1-10, 11-20, 21-30, 31-40, and 41-50, is covered. The price identification task was incentivized: One trial was randomly selected to determine the bonus payment and subjects received £0.5 if they identified all five prices in this trial correctly. We thereby aim to shift attention away from individual information brackets towards the entire price path. If the attention manipulation is successful, beliefs will move closer to those in the broad information treatment of the baseline experiment.

Figure 5 plots subjects' average estimation error for our attention manipulation (Experiment 2; in yellow) and compares it to the narrow (blue) and broad (red) information treatment from the baseline experiment.

Figure 5: Experiment 2: Estimation Error



Note: This figure plots the average estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) in Experiment 2 in comparison to the narrow and broad information treatment of Experiment 1 for negative respectively positive price paths.

Consistent with our conjecture, we find that estimation errors in the attention manipulation are significantly lower than those in the narrow treatment of the baseline experiment ($p < 0.01$ for both negative and positive paths). Additionally, we find that estimation errors in the attention manipulation (6.5 and 6.8 for negative and positive price paths, respectively) are now statistically indistinguishable from estimation errors in the broad treatment of our baseline experiment (6.3 and 6.1, for negative and positive price paths, respectively). As such, exogenously inducing attention successfully closes the gap between the narrow and the broad information treatment observed in the baseline experiment. Table B2 in Appendix B confirms this in a regression setting.

Next, we investigate whether the attention manipulation affects how subjects weight recent versus distant information when forming their beliefs. Specifically, we again run the augmented Greenwood and Shleifer (2014) regression model but estimate an additional δ coefficient for subjects in the attention manipulation:

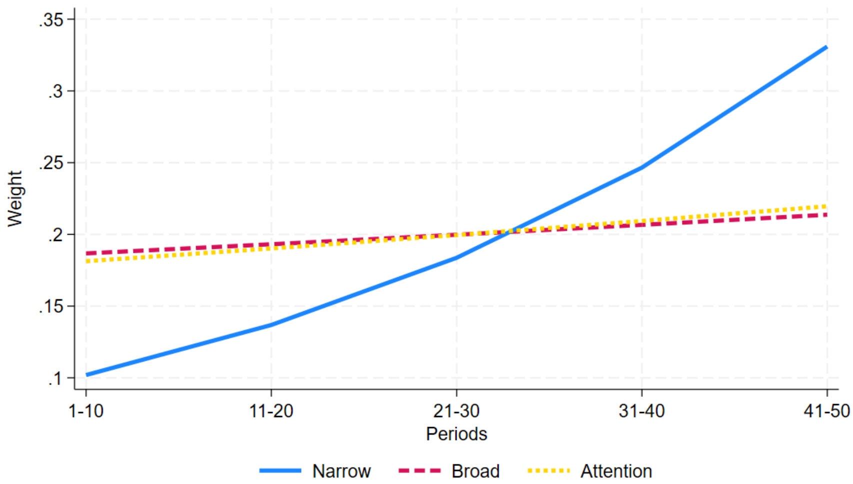
$$p_i = a + b \cdot \sum_{j=0}^{k} \omega_j \Delta Price_{t-j} + u_t, \tag{10}$$

with

$$\omega_j = \frac{(\delta_{broad} + \delta_{narrow} \cdot Narrow + \delta_{attention} \cdot Attention)^j}{\sum_{l=0}^{k}(\delta_{broad} + \delta_{narrow} \cdot Narrow + \delta_{attention} \cdot Attention)^l}.$$

Table B3 in Appendix B presents the regression results. They are illustrated in Figure 6. As subjects in the broad information treatment of our baseline experiment, subjects in the attention manipulation place almost equal weight on all information brackets. Compared to subjects in the narrow information treatment of our baseline experiment, however, subjects in the attention manipulation underweight recent information and overweight distant information. Attention thus directly affects how information is incorporated into subjects' beliefs.

Figure 6: Experiment 2: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 10 in Experiment 2 compared to the narrow and broad information treatment of Experiment 1.

Finally, we use subjects' answers to the memory questions to validate that our manipulation indeed affects attention and ultimately memory. The results of this exercise are reported in Table 5. Panel A shows that the fraction of correct answers in the attention manipulation is higher than in the baseline for each of the 3 questions asked in the attention manipulation.[11] In comparison to the broad information treatment, the differences in fractions are small and not significant, but in comparison to the narrow information treatment the differences are highly significant. The attention manipulation increases the fraction of correct answers relative

[11] In comparison to the baseline experiment, the recall task in the attention manipulation only comprises the questions regarding the number of positive and negative price changes, as well as the final price in period 50.

to the narrow information treatment by at least 50%, resulting in a level of correct answers similar to the one in the broad information treatment. Panel B confirms this impression using the number of correctly answered questions. We conclude that our attention manipulation was successful in shifting subjects' attention towards the entire price path, resulting in a better ability to recall the provided information.

Table 5: Experiment 2: Memory

Panel A: Fraction			in %		
	Attention	Broad	Difference	Narrow	Difference
increases	23.58	19.81	3.77	13.57	10.01***
			(1.04)		(3.77)
decreases	23.58	19.81	3.77	13.77	9.81***
			(1.04)		(3.68)
final price	24.15	25.00	-0.85	16.37	7.78***
			(-0.23)		(2.82)
Panel B: Number					
	Attention	Broad	Difference	Narrow	Difference
memory score	0.71	0.65	0.07	0.44	0.28***
			(0.58)		(4.68)

Note: This table displays answers to the memory questions of Experiment 2 in comparison to the broad respectively narrow information treatment of Experiment 1. Panel A displays the fraction of subjects who answered correctly per question and Panel B displays the memory score out of all 3 questions asked in Experiment 2. In both panels, Mann-Whitney tests are used to test for differences in means; the corresponding z-scores are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

5 Applications

In this section, we explore how information partitioning influence belief formation in two applications. We first study an application to financial markets in which recent information carries greater value than older information. Second, we study an application to consumer choice, in which customers judge the quality of a good based on online reviews.

5.1 Financial Markets

One of the basic premises in financial economics is that market prices are very informative about the fundamentals of traded assets (Goldstein, 2023). However, not all available (recent and past) prices are equally informative. If markets are efficient, stock prices should rapidly incorporate all value-relevant signals and thus quickly become stale information. This distinctive feature is different from the environment studied in our baseline experiment, in which all information is equally important for the formation of beliefs. Our results so far show that individuals who process information in narrow brackets overweight recent information, leading to less accurate beliefs compared to individuals who process information in broad brackets. However, if recent information is also more informative, processing information in narrow brackets may result in better calibrated beliefs. We test this conjecture in a third experiment (Experiment 3; $N = 998$).

The third experiment builds directly on our baseline design. To incorporate the notion that recent information is more important, we change the underlying data-generating process from which stock price movements are generated. Specifically, the probability of positive and negative price changes ($+/-10$) now depends on the underlying state of the asset. At first, the asset can be in a good state or bad state with equal probability. If the asset is in the good state, the probability of a positive return is 70% (and negative return 30%). If the asset is in the bad state, the probability of a positive return is 30% (and negative return 70%). The state of the asset remains fixed for 10 periods (out of 50). After every 10th period the state of the asset changes with 25% probability and remains the same with 75% probability. As such, price changes are governed by a Markov chain as in Frydman et al. (2014) and Charles et al. (2024). A direct consequence of this Markov chain is that a Bayesian agent would overweight recent signals as they are more diagnostic of the current state compared to older signals. The data-generating process is known to the subjects, and their understanding of it is ensured by comprehension checks. In addition to employing a price process that represents financial markets, we conduct the experiment on a sample of more sophisticated investors who report to regularly participate in the stock market.

We start by investigating subjects' estimation error. Table 6 displays estimates of Equation 8 for the whole sample (Columns (1) and (2)), as well as for positive and negative price paths separately (Columns (3) and (4), respectively). The coefficient of $Narrow$ is highly statistically significant and of similar magnitude across all specifications, confirming that information partitioning also affects beliefs in a setting with a different data-generating process and in a sample of more financially sophisticated investors. Note that the coefficient is negative, indicating that subjects in the narrow treatment form beliefs that are closer to the Bayesian benchmark. Although this seems opposite to our previous findings, it is behaviorally consistent with a data-generating process in which recent information is more informative about the underlying state. This conjecture rests on the assumption that individuals in the

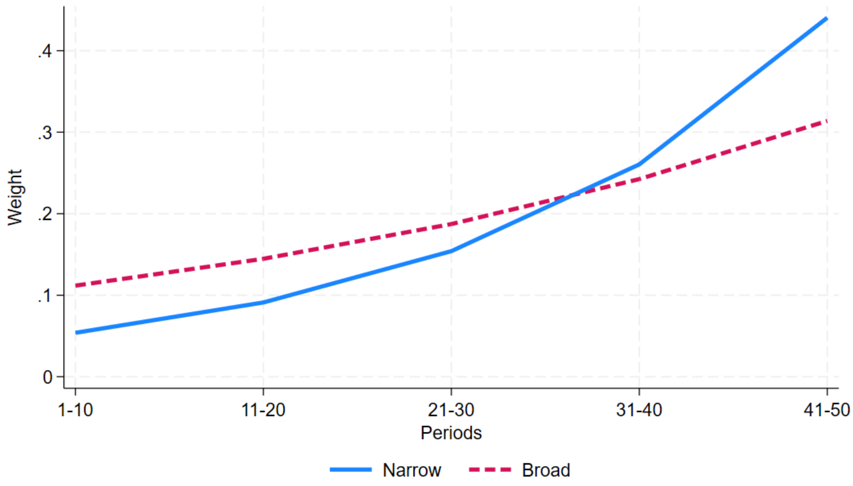
Table 6: Finance Application: Estimation Error

	Overall	Overall	Negative	Positive
bayes	-0.01	-0.01	0.18^{***}	-0.18^{***}
	(0.01)	(0.01)	(0.02)	(0.02)
narrow	-3.80^{***}	-3.69^{***}	-4.23^{***}	-3.07^{***}
	(0.91)	(0.91)	(1.17)	(1.07)
Controls	No	Yes	Yes	Yes
N	3,024	3,000	1,704	1,296
R^2	0.01	0.01	0.12	0.16

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, the narrow information dummy and control variables as independent variables for the Financa Application. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

narrow treatment put more weight on recent information in their belief formation than those in the broad treatment, as documented in our main experiment. To ensure that this is still the case, we estimate Equation 9 for the finance application. Coefficient estimates are displayed in Columns (1) to (3) of Table B4 in Appendix B. The results are illustrated in Figure 7. In line with the new data generating process, subjects in either treatment place more weight on recent information relative to distant information. We nevertheless find that individuals who process information in narrow brackets put more weight on recent information than those who process information in broad brackets. The observed behavior across experiments is thus identical. However, the underlying data-generating process ultimately decides whether the behavior leads to better or worse calibrated beliefs.

Figure 7: Finance Application: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 9 in the Finance Application.

5.2 Consumer Choice

In many online market places, consumers can learn about the quality of a good or service by reading reviews from other customers. Similarly to the information environment studied in our baseline setting, online market places typically place more equal weight on all available reviews as long as the product remains the same and is not replaced by an updated version. In contrast to the settings studied so far, these reviews are often more qualitative in nature, as customers write about their experiences and opinions. To investigate whether the partitioning of (qualitative) information into smaller or larger brackets affects consumer judgments, we conduct a fourth experiment (Experiment 4; $N = 996$).

In the experiment, subjects learn about the quality of a fictional smartphone. We chose smartphones as they are generic products with a stable quality in between new releases. To assess the quality, subjects observe a total of 15 product ratings. Each rating consists of an overall score (displayed as a 1 to 5 star rating), a short summary statement of the rating, as well as one good and one bad aspect of the phone.[12] After observing the ratings, subjects assess the quality of the phone on a scale from 1 (very bad) to 7 (very good) in 0.1 increments.

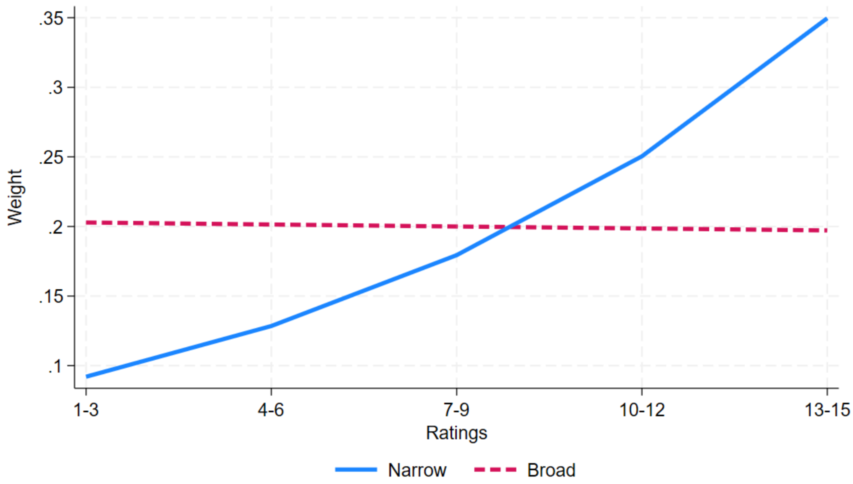
[12] All product ratings displayed in the study are AI generated without reference to a brand. Each rating has been manually evaluated to ensure its appropriateness and whether the wording is comprehensible. Screenshots of the display format are contained in Figures C11 and C12 in Appendix C.

The question wording was inspired by prior research on product ratings (e.g., Floyd et al., 2014; Ziegele and Weber, 2015). To establish a reasonable benchmark for quality, we follow Bhatt et al. (2015) and classify products for which the majority of the ratings display 4 and 5 stars as "good" products and products for which the majority of the ratings display 1 and 2 stars as "bad" products. We drew multiple sets of ratings such that the cumulative signals imply a good product in some sets and a bad product in other sets. As in our baseline experiment, we exogenously vary the frequency at which ratings are processed. In the narrow information treatment, we elicit subjects' quality assessment after every third rating and as such 5 times in total. Although subjects observe ratings in batches of 3, all previous ratings are displayed jointly with the new batch of ratings to avoid memory concerns and make the comparison as clean as possible. Unlike in our previous experiments, the ratings do not have an inherent time dimension, and all 3 ratings of a batch are displayed simultaneously. In the broad information treatment, we elicit subjects' quality assessment only once after all 15 ratings have been observed.

Even in this rather simplistic setting with qualitative information signals, we find that partitioning information in smaller or larger brackets causally affects judgment. Specifically, we find that processing reviews in larger brackets (broad treatment) leads to more optimistic product ratings for good products (4.30 vs. 3.99, for broad and narrow, respectively; $p < 0.01$). For bad products, we observe a similar pattern, although less pronounced and not statistically significant (3.35 vs. 3.24, for broad and narrow respectively; $p > 0.1$), likely because consumers have much more differentiated opinions for goods with positive rather than negative evaluations in online commerce settings (Bhatt et al., 2015).

Since it is difficult to judge whether subjective ratings in the narrow or broad treatment better reflect subjects' true opinion in absence of a clear benchmark, we focus on how information is weighted in their final assessment. This allows us to infer whether the behavior is internally consistent with the behavior documented in our previous settings. To do so, we again compare how information early and late in the sequence affects final judgments between the broad and narrow treatment. Coefficient estimates of Equation 9 are displayed in Columns (4) to (6) of Table B4 in Appendix B. Figure 8 illustrates the findings. They

Figure 8: Consumer Choice Application: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 9 in the Consumer Choice Application.

confirm our previous conjecture. In line with the idea that product rating do not have an inherent time dimension we observe that subjects in the broad treatment place equal weight on all 15 ratings. Importantly, subjects in the narrow treatment still put more weight on recently observed information than subjects in the broad treatment. Overall, this shows that our results are highly consistent across information types (qualitative vs. quantitative), visual presentation (price chart vs. list), settings (financial markets vs. online market places), and data-generating processes (equal-weighted information signals vs. Markov chain).

6 Alternative Explanations

6.1 Number of Trials

Subjects in the narrow information treatment provide estimates more frequently and thus spend more time and effort on each asset than subjects in the broad information treatment. To avoid cognitive fatigue and enable a fair comparison across treatments, we confront subjects in the narrow information treatment with only four instead of eight trials. While our experimental design does not provide feedback after each trial, subjects might still become more familiar with the setting over the course of these trials. We therefore check whether the different number of trials across treatments affects our results. We distinguish between the first and the second half of trials per treatment. For subjects in the narrow (broad) information treatment, estimates of the first 2 (4) trials are in the first half, and estimates of the last 2 (4) trials are in the second half. Our findings remain in each half. Even in the second half of our experiment, subjects in the narrow information treatment exhibit greater estimation errors and put more weight on recent rather than distant observations relative to subjects in the broad information treatment (See Table B5 and B6 in Appendix B). Also note that in the finance application, subjects in the narrow treatment provide more accurate estimates than subjects in the broad treatment albeit facing a smaller number of trials. The fact that subjects in the broad information treatment face more trials than subjects in the narrow information treatment cannot explain the information partitioning effect.

6.2 Attentiveness

Another potential concern is that some subjects are overburdened with the experiment and as a result rush through it without being attentive and making an effort. Since subjects in the narrow treatment have to provide more estimates per trial, they might become careless more easily. First, note that the analysis in the previous subsection confirms our results for the first half of trials, i.e., when attentiveness in both treatments is still high. Second, we investigate participants' total working time as a proxy for effort. If carelessness is at play, we should observe that those participants who completed the experiment the quickest are driving our results. We follow Enke and Graeber (2023) and define subjects who are in the bottom quintile of the total working time distribution as speeders. Table B7 and B8 in Appendix B report results of our main analysis excluding these speeders. If anything, results for attentive subjects are even stronger than for the entire sample. The information partitioning effect is not driven by inattentiveness.

6.3 Statistical Skill

The question remains whether our findings are transferable to more experienced and sophisticated subjects. Note that subjects in our experiments are required to answer three comprehension questions correctly before they can proceed to the actual task to ensure their understanding of the underlying setting (See Appendix C). Within this sample, we further distinguish between subjects with low respectively high self-reported statistical skill.[13] While higher statistical skill is associated with lower estimation errors in general, the effect of information partitioning on estimation errors is even stronger among high-skilled subjects (Table B9). Processing information at higher frequency also leads subjects with high skill to overweight recent information and underweight distant information relative to processing information at lower frequency (Table B10). The information partitioning effect thus applies to both naive and sophisticated subjects.

6.4 Recency of Information Arrival

Lastly, we investigate the role of intermediate belief updating in the narrow treatment. Specifically, subjects in our narrow information treatments not only observe the information at a higher frequency, but also provide an intermediate estimate after each information bracket. In standard models of belief updating – like the one in our framework – the influence of observing information at different frequencies and the updating of beliefs cannot be separated, as individuals are usually assumed to update whenever new information arrives. Our experiments were designed to reflect this property and to ensure that each subject processes the provided information in the narrow treatment.

First note that subjects in the narrow treatment of our baseline experiment apply almost equal weight to all available information when they provide their first estimate in period 10. When there are no intermediate belief updates, subjects in the narrow treatment seem to behave as subjects in the broad treatment. To demonstrate this more rigorously, we ran a third treatment in our finance application in which subjects observe information in narrow brackets (i.e., the price-line chart slowly builds over time) but only provide one final estimate. We call this the *mixed* treatment. This allows us to disentangle recency effects in information arrival from the frequency at which information is processed. Subjects in both the narrow and the mixed treatment observe information at the same frequency (one signal every 0.5

[13] Subjects who reported above (below or equal to) median statistical skill belong to the high (low) skill subsample.

seconds), but process information at different frequencies (every 10 periods vs. once in period 50). Any difference in beliefs is thus attributable to the frequency at which information is processed. Results for the information weights are displayed in Figure B3 in Appendix B. As can be inferred, subjects in the mixed treatment behave exactly as subjects in the broad treatment. Relative to subjects in the narrow treatment, subjects in the mixed treatment place more equal weight on all information. Our main finding thus cannot be explained by the recency of information arrival, but relies on information being *processed* at different frequencies.

7 Conclusion

In this paper, we study the influence of information partitioning on learning and beliefs. We show that partitioning information influences how individuals incorporate such information into their expectations. Processing information in smaller amounts more frequently (narrow information) causes individuals to overweight recent information and to underweight distant information. Similar behavior cannot be observed if information is processed less frequently but in larger batches (broad information), where individuals put equal weight on recent and distant information. We further show that depending on the underlying data-generating process, the overweighting of recent information can lead to less or more accurate judgments. In exploring the mechanism, we demonstrate that partitioning information into narrower brackets channels attention towards isolated information signals rather than the joint set of information. This heightened attention to small pieces of information not only leads to overextrapolation from recent signals, but also to significantly worse recall of the encountered information, consistent with the model of diagnostic expectations (Bordalo et al., 2016; Bordalo et al., 2018; Bordalo et al., 2019).

Our results imply that breaking information into smaller or larger partitions can be a powerful tool to alter individuals' expectations with applications in diverse fields. For instance, firms or information providers such as financial advisors often choose whether to disclose information regarding company performance or product ratings in narrower or broader brackets. Such choices could either willingly or unwillingly manipulate their clients' judgments. On a broader scale, one may argue that narrow bracketing enables many well-documented errors in probabilistic reasoning. For instance, individuals' belief in the law of small numbers – i.e., the belief that small random samples are highly representative of their underlying population (Tversky and Kahneman, 1971) – is most severe if information is processed in small brackets, but would have a smaller impact on judgments if information is presented (and processed) in larger brackets. Similarly, base-rate neglect – the fact that people on average under-use prior information (Kahneman and Tversky, 1973) – causes individuals to "jump to conclusions" when presented with small information samples but leads to persistent uncertainty when presented with larger samples (Benjamin, 2019). As such, the implications of base-rate neglect for belief updating ultimately depend on whether information is processed in narrow or broad brackets.

References

Augenblick, N., Lazarus, E., Thaler, M., 2025. Overinference from weak signals and underinference from strong signals. *The Quarterly Journal of Economics* 140, 335–401.

Ba, C., Bohren, J. A., Imas, A., 2024. Over-and underreaction to information. *Available at SSRN 4274617* .

Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-capm: An extrapolative capital asset pricing model. *Journal of Financial Economics* 115, 1–24.

Bénabou, R., Tirole, J., 2002. Self-confidence and personal motivation. *Quarterly Journal of Economics* 117, 871–985.

Benartzi, S., Thaler, R. H., 1999. Risk aversion or myopia? choices in repeated gambles and retirement investments. *Management Science* 45, 364–381.

Benjamin, D. J., 2019. Errors in probabilistic reasoning and judgment biases. In: *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, pp. 69–186.

Benjamin, D. J., Rabin, M., Raymond, C., 2016. A model of nonbelief in the law of large numbers. *Journal of the European Economic Association* 14, 515–544.

Bhatt, A., Patel, A., Chheda, H., Gawande, K., 2015. Amazon review classification and sentiment analysis. *International Journal of Computer Science and Information Technologies* 6, 5107–5110.

Bohren, J. A., Hascher, J., Imas, A., Ungeheuer, M., Weber, M., 2024. A cognitive foundation for perceiving uncertainty. Tech. rep., National Bureau of Economic Research.

Bordalo, P., Burro, G., Coffman, K., Gennaioli, N., Shleifer, A., 2025. Imagining the future: memory, simulation, and beliefs. *Review of Economic Studies* 92, 1532–1563.

Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2016. Stereotypes. *The Quarterly Journal of Economics* 131, 1753–1794.

Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2019. Diagnostic expectations and stock returns. *The Journal of Finance* 74, 2839–2874.

Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Overreaction in macroeconomic expectations. *American Economic Review* 110, 2748–2782.

Bordalo, P., Gennaioli, N., Shleifer, A., 2018. Diagnostic expectations and credit cycles. *The Journal of Finance* 73, 199–227.

Brunnermeier, M. K., Parker, J. A., 2005. Optimal expectations. *American Economic Review* 95, 1092–1118.

Camerer, C., Hua Ho, T., 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.

Charles, C., Frydman, C., Kilic, M., 2024. Insensitive investors. *The Journal of Finance* 79, 2473–2503.

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

Chew, S. H., Huang, W., Zhao, X., 2020. Motivated false memory. *Journal of Political Economy* 128, 3913–3939.

Cripps, M. W., 2018. Divisible updating. *Manuscript, UCL* .

Danz, D., Vesterlund, L., Wilson, A. J., 2022. Belief elicitation and behavioral incentive compatibility. *American Economic Review* 112, 2851–2883.

Edwards, W., Phillips, L. D., 1964. *Man as tranducer for probabilities in bayesian command and control systems* .

Ellis, A., Freeman, D. J., 2024. Revealing choice bracketing. *American Economic Review* 114, 2668–2700.

Enke, B., Graeber, T., 2023. Cognitive uncertainty. *The Quarterly Journal of Economics* 138, 2021–2067.

Enke, B., Schwerter, F., Zimmermann, F., 2020. Associative memory and belief formation. Tech. rep., National Bureau of Economic Research.

Enke, B., Schwerter, F., Zimmermann, F., 2024. Associative memory, beliefs and market interactions. *Journal of Financial Economics* 157, 103853.

Enke, B., Zimmermann, F., 2019. Correlation neglect in belief formation. *The Review of Economic Studies* 86, 313–332.

Fischbacher, U., Hoffmann, G., Schudy, S., 2017. The causal effect of stop-loss and take-gain orders on the disposition effect. *The Review of Financial Studies* 30, 2110–2129.

Fischhoff, B., Bar-Hillel, M., 1984. Diagnosticity and the base-rate effect. *Memory & Cognition* 12, 402–410.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., Freling, T., 2014. How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing* 90, 217–232.

Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., Rangel, A., 2014. Insensitive investors. *The Journal of Finance* 69, 907–946.

Fudenberg, D., Levine, D. K., et al., 2014. Learning with recency bias. *Proceedings of the National Academy of Sciences* 111, 10826–10829.

Gabaix, X., 2019. Behavioral inattention. In: *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, pp. 261–343.

Gagnon-Bartsch, T., Bushong, B., 2022. Learning with misattribution of reference dependence. *Journal of Economic Theory* 203, 105473.

Gennaioli, N., Shleifer, A., 2010. What comes to mind. *The Quarterly Journal of Economics* 125, 1399–1433.

Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics* 112, 631–645.

Gödker, K., Jiao, P., Smeets, P., 2025. Investor memory. *Review of Financial Studies* 38, 1595–1640.

Goldstein, I., 2023. Information in financial markets and its real effects. *Review of Finance* 27, 1–32.

Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *The Review of Financial Studies* 27, 714–746.

Hartzmark, S. M., Hirshman, S. D., Imas, A., 2021. Ownership, learning, and beliefs. *The Quarterly Journal of Economics* 136, 1665–1717.

He, X. D., Xiao, D., 2017. Processing consistency in non-bayesian inference. *Journal of Mathematical Economics* 70, 90–104.

Hirshleifer, D., Li, J., Yu, J., 2015. Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics* 76, 87–106.

Jiang, Z., Liu, H., Peng, C., Yan, H., 2024. Investor memory and biased beliefs: Evidence from the field. Tech. rep., National Bureau of Economic Research.

Jiang, Z., Liu, H., Peng, C., Yan, H., 2025. Investor memory and biased beliefs: Evidence from the field. *Available at Western Finance Portal 875336* .

Kahana, M. J., 2012. Foundations of human memory. OUP USA.

Kahneman, D., Tversky, A., 1972. Subjective probability: A judgment of representativeness. *Cognitive psychology* 3, 430–454.

Kahneman, D., Tversky, A., 1973. On the psychology of prediction. *Psychological Review* 80, 237.

Kieren, P., Weber, M., 2025. Expectation formation under uninformative signals. *Management Science* 71, 5123–5141.

Kőszegi, B., Matějka, F., 2020. Choice simplification: A theory of mental budgeting and naive diversification. *The Quarterly Journal of Economics* 135, 1153–1207.

Kunda, Z., 1990. The case for motivated reasoning. *Psychological bulletin* 108, 480.

Malmendier, U., Pouzo, D., Vanasco, V., 2020. Investor experiences and financial market dynamics. *Journal of Financial Economics* 136, 597–622.

Moore, D. A., Healy, P. J., 2008. The trouble with overconfidence. *Psychological review* 115, 502.

Mrkva, K., Van Boven, L., 2017. Attentional accounting: Voluntary spatial attention increases budget category prioritization. *Journal of Experimental Psychology: General* 146, 1296.

Mullainathan, S., 2002. A memory-based model of bounded rationality. *The Quarterly Journal of Economics* 117, 735–774.

Phillips, L. D., Edwards, W., 1966. Conservatism in a simple probability inference task. *Journal of Experimental Psychology* 72, 346.

Read, D., Loewenstein, G., Rabin, M., 1999. Choice bracketing. *Journal of Risk and Uncertainty* 19, 171–197.

Ross, L., 1977. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In: *Advances in Experimental Social Psychology*, Elsevier, vol. 10, pp. 173–220.

Schwartzstein, J., 2014. Selective attention and learning. *Journal of the European Economic Association* 12, 1423–1452.

Tversky, A., Kahneman, D., 1971. Belief in the law of small numbers. *Psychological bulletin* 76, 105.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. *Science* 185, 1124–1131.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 453–458.

Ungeheuer, M., Weber, M., 2021. The perception of dependence, investment decisions, and stock prices. *The Journal of Finance* 76, 797–844.

Verghese, P., 2001. Visual search and attention: A signal detection theory approach. *Neuron* 31, 523–535.

Wachter, J. A., Kahana, M. J., 2024. A retrieved-context theory of financial decisions. *The Quarterly Journal of Economics* 139, 1095–1147.

Ziegele, M., Weber, M., 2015. Example, please! comparing the effects of single customer reviews and aggregate review scores on online shoppers' product evaluations. *Journal of Consumer Behaviour* 14, 103–114.

Zimmermann, F., 2020. The dynamics of motivated beliefs. *American Economic Review* 110, 337–363.

A Proofs and Derivations

A.1 Proof of Proposition 1

Proposition 1: For $\lambda = 1$, partitioning information into smaller or larger brackets does not influence posterior beliefs ($\pi_t^{j,\text{narrow}} = \pi_t^{j,\text{broad}}$).

Proof: We first show that for $\lambda = 1$, $\mu_t^{j,\text{narrow}} = \mu_t^{j,\text{broad}} = \mu_t^{j,\text{Bayes}}$.
First, notice that for $\lambda = 1$, we obtain $\pi_t^j = \mu_t^j$ for any t. We first show that $\mu_t^{j,\text{narrow}} = \mu_t^{j,\text{Bayes}}$. Afterwards, we show that $\mu_t^{j,\text{broad}} = \mu_t^{j,\text{Bayes}}$, from which $\mu_t^{j,\text{narrow}} = \mu_t^{j,\text{broad}}$ follows.

Narrow Information Processing:

We start by rewriting the model in the posterior-odds form:

$$\frac{\mu_t^{G,\text{narrow}}}{\mu_t^{B,\text{narrow}}} = \frac{P(s_t \mid \theta_G)}{P(s_t \mid \theta_B)} \cdot \frac{P(S_{t-1} \mid \theta_G)}{P(S_{t-1} \mid \theta_B)}, \tag{11}$$

where $P(s_t \mid \theta_j)$ is the probability of observing signal s_t conditional on the type $j \in \{G, B\}$, and $P(S_{t-1} \mid \theta_G)$ is the probability of observing the full signal history S_{t-1} until period $t-1$ conditional on type $j \in \{G, B\}$.
Next, notice that $P(S_{t-1} \mid \theta_j)$ can be rewritten as $P(s_{t-1} \mid \theta_j)P(S_{t-2} \mid \theta_j)$ as the signal process is i.i.d. By continuously substituting, Equation (11) can be rewritten as:

$$\frac{\mu_t^{G,\text{narrow}}}{\mu_t^{B,\text{narrow}}} = \frac{P(S_t \mid \theta_G)}{P(S_t \mid \theta_B)} \cdot \frac{\pi_0^G}{\pi_0^B}, \tag{12}$$

where π_0^j represents the agent's prior belief about the distribution of good and bad types. This corresponds to the standard Bayesian updating process.

Broad Information Processing:

We again rewrite the model in the posterior-odds form:

$$\frac{\mu_t^{G,\text{broad}}}{\mu_t^{B,\text{broad}}} = \frac{P(s_{t:n} \mid \theta_G)}{P(s_{t:n} \mid \theta_B)} \cdot \frac{P(S_{t-n} \mid \theta_G)}{P(S_{t-n} \mid \theta_B)}, \tag{13}$$

where $P(s_{t:n} \mid \theta_j)$ is the probability of observing signals from a batch of size n in period t conditional on type $j \in \{G, B\}$, and $P(S_{t-n} \mid \theta_G)$ is the probability of observing the full signal history S_{t-n} until period $t-n$ conditional on type $j \in \{G, B\}$.

Notice that $P(s_{t:n} \mid \theta_j)$ can be rewritten as $\prod_{i=0}^{n-1} P(s_{t-i} \mid \theta_j)$ as the signal process is i.i.d. Additionally, $P(S_{t-n} \mid \theta_j)$ can be rewritten as $P(s_{t-n} \mid \theta_j)P(S_{t-n-1} \mid \theta_j)$. By continuously substituting, Equation (13) can be rewritten as:

$$\frac{\mu_t^{G,\text{broad}}}{\mu_t^{B,\text{broad}}} = \frac{P(S_t \mid \theta_G)}{P(S_t \mid \theta_B)} \cdot \frac{\pi_0^G}{\pi_0^B}, \tag{14}$$

which is the same as $\frac{\mu_t^{G,\text{narrow}}}{\mu_t^{B,\text{narrow}}}$.

Since for $\lambda = 1$, we obtain $\pi_t^j = \mu_t^j$, it follows that $\pi_t^{j,\text{narrow}} = \pi_t^{j,\text{broad}}$.

A.2 Proof of Proposition 2

Proposition 2: For $\lambda \neq 1$, partitioning information into smaller or larger brackets generally results in different posterior beliefs ($\pi_t^{j,\text{narrow}} \neq \pi_t^{j,\text{broad}}$).

Proof: We first show that for narrow information processing the order of signals matters, whereas for broad information processing it does not. We then show that the two processes lead to generally different posterior beliefs. Note that for $t = 1$ the updating processes are identical as it has to follow that $n = 1$, too. We thus show that they are different for $t > 1$.

Order of signals: Let σ be a permutation of the set $\Sigma = \{1, 2, \ldots, t\}$, and define the permuted signal sequence as $S_t' = (s_{\sigma(1)}, \ldots, s_{\sigma(t)})$ with $S_t' \neq S_t$.

Narrow Information Processing:

When processing one signal at a time ($n = 1$), by iterating:

$$\pi_k^{G,\text{narrow}}(S_k) = (1-\lambda)\pi_{k-1}^{G,\text{narrow}} + \lambda\mu_k^{G,\text{narrow}} \quad \text{for } k = 1, \ldots, t \tag{15}$$

one shows by recursion that

$$\pi_t^{G,\text{narrow}}(S_t) = (1-\lambda)^t \pi_0^G + \lambda \sum_{i=1}^{t} (1-\lambda)^{t-i} \mu_i^{G,\text{narrow}} \tag{16}$$

This is a similar algebraic expansion as used in Equation 12 for $\lambda = 1$, but now with an extra $(1-\lambda)$ dampening at each iteration due to $\lambda \neq 1$.

Consider the permuted belief as:

$$\pi_t^{G,\text{narrow}}(S_t') = (1-\lambda)^t \pi_0^G + \lambda \sum_{i=1}^{t} (1-\lambda)^{t-i} \mu_{\sigma(i)}^{G,\text{narrow}} \tag{17}$$

Then, for any $\lambda \neq 1$, there exist signal sequences S_t and S_t', which are permutations of one another (i.e., have the same number of g and b signals), such that

$$\pi_t^{G,\text{narrow}}(S_t) \neq \pi_t^{G,\text{narrow}}(S_t') \tag{18}$$

That is, the narrow belief update is not invariant to the order of signals. To show this, we proceed by constructing a simple counterexample for $t = 2$. Consider the following two sequences of signals $S_2 = (g, b)$ and $S'_{\sigma(2)} = (b, g)$.

Case 1: $S_2 = (g, b)$

The agent's first period belief is

$$\pi_1^{G,\text{narrow}}(g) = (1-\lambda)\pi_0^G + \lambda\mu_1^{G,\text{narrow}}(g; \pi_0^G), \tag{19}$$

where $\mu_1^{G,\text{narrow}}(g; \pi_0^G)$ denotes the Bayesian posterior for observing signal g conditional on the prior π_0^G. Then, the agent's second period belief is

$$\pi_2^{G,\text{narrow}}(g, b) = (1-\lambda)\pi_1^{G,\text{narrow}}(g) + \lambda\mu_1^{G,\text{narrow}}(b; \pi_1^{G,\text{narrow}}(g)), \tag{20}$$

where $\mu_1^{G,\text{narrow}}(b; \pi_1^{G,\text{narrow}}(g))$ denotes the Bayesian posterior for observing signal b conditional on the prior $\pi_1^{G,\text{narrow}}(g)$.

Case 2: $S'_{\sigma(2)} = (b, g)$

The agent's first period belief is

$$\pi_1^{G,\text{narrow}}(b) = (1-\lambda)\pi_0^G + \lambda\mu_1^{G,\text{narrow}}(b; \pi_0^G), \tag{21}$$

where $\mu_1^{G,\text{narrow}}(b; \pi_0^G)$ denotes the Bayesian posterior for observing signal b conditional on the prior π_0^G. The second period belief is

$$\pi_2^{G,\text{narrow}}(b, g) = (1-\lambda)\pi_1^{G,\text{narrow}}(b) + \lambda\mu_1^{G,\text{narrow}}(g; \pi_1^{G,\text{narrow}}(b)), \tag{22}$$

where $\mu_1^{G,\text{narrow}}(g; \pi_1^{G,\text{narrow}}(b))$ denotes the Bayesian posterior for observing signal g conditional on the prior $\pi_1^{G,\text{narrow}}(b)$.

Comparing both second period beliefs, note that $\mu_1^{G,\text{narrow}}(g; \pi_0^G) > \pi_0^G > \mu_1^{G,\text{narrow}}(b; \pi_0^G)$ given that $0 < \theta_B < \theta_G < 1$, i.e., signal g increases posteriors about type G, while signal b decreases them. Given that the function is increasing in $\mu_1^{(G,\text{narrow})}$, it follows that $\pi_1^{G,\text{narrow}}(g) > \pi_0^G > \pi_1^{G,\text{narrow}}(b)$. But then, because the posterior depends nonlinearly on the prior for $\lambda \neq 1$, we generally obtain that

$$\mu_1^{G,\text{narrow}}(b;\pi_1^{G,\text{narrow}}(g)) \neq \mu_1^{G,\text{narrow}}(g;\pi_1^{G,\text{narrow}}(b)), \tag{23}$$

and therefore

$$\pi_2^{G,\text{narrow}}(g,b) \neq \pi_2^{G,\text{narrow}}(b,g). \tag{24}$$

As such, the narrow belief updating for $\lambda \neq 1$ is non-commutative.

Broad Information Processing:

When processing the information in batches of size $n > 1$, the updating rule is:

$$\pi_t^{G,\text{broad}}(S_t) = (1-\lambda)\pi_{t-n}^{G,\text{broad}} + \lambda\mu_t^{G,\text{broad}} \tag{25}$$

First, consider the agent's belief after observing the first batch of signals ($t = n$), which simplifies the updating rule to

$$\pi_t^{G,\text{broad}}(S_t) = (1-\lambda)\pi_0^{G} + \lambda\mu_t^{G,\text{broad}} \tag{26}$$

Since $\lambda \neq 1$, the ordering of signals within a batch can only matter if and only if it affects the posterior implied by the most recent batch of signals $\mu_t^{G,\text{broad}}$. However, because signals are i.i.d. by assumption, the ordering within a batch of signals is irrelevant:

$$P(S_t|\theta_j) = \prod_{i=1}^{t} P(s_i|\theta_j) = \prod_{i=1}^{t} P(s_{\sigma(i)}|\theta_j) = P(S_{\sigma(t)}|\theta_j) \tag{27}$$

From Equation (27) immediately follows that the posterior implied by the signal $\mu_t^{G,\text{broad}}$ is independent of the ordering of signals. Thus, by repeated substitution one can show that

$$\pi_t^{G,\text{broad}}(S_t) = \pi_t^{G,\text{broad}}(S_{\sigma(t)}) \tag{28}$$

holds for any $t > n$, as long as the ordering across batches does not change. Note that when the ordering across batches changes too, results for the broad information processing converge to narrow information processing for $n \to 1$, and diverge for $n \to T$.

Determining the Difference:

$$\Delta = \pi_t^{G,\text{narrow}} - \pi_t^{G,\text{broad}} = \lambda \left[\sum_{i=1}^{t} (1-\lambda)^{t-i} \mu_i^{G,\text{narrow}} - \mu_t^{G,\text{broad}} \right] + [(1-\lambda)^t - (1-\lambda)]\pi_0^G \quad (29)$$

For this difference to be zero (and the posteriors between the two processes to be zero) we thus require

$$\lambda \left[\sum_{i=1}^{t} (1-\lambda)^{t-i} \mu_i^{G,\text{narrow}} - \mu_t^{G,\text{broad}} \right] = [(1-\lambda) - (1-\lambda)^t]\pi_0^G \quad (30)$$

The RHS is non-zero for $\lambda \neq 1$, but only depends on λ and the scalar π_0^G. It is independent of the signal ordering and constant for a given λ and prior π_0^G. The LHS, however, depends on the signal ordering. It thus cannot generally equal the RHS.

It follows that $\pi_t^{j,\text{narrow}}$ is generally not equal to $\pi_t^{j,\text{broad}}$.

A.3 Proof of Proposition 3

Proposition 3: For $\lambda \neq 1$, there is no simple comparative static that describes whether narrow or broad information processing makes individuals consistently more optimistic or pessimistic than the other.

Proof: To show this, we construct two counterexamples for $t = 3$. In the first signal sequence, narrow updating leads to a lower posterior belief than broad updating $\pi_t^{G,\text{narrow}} < \pi_t^{G,\text{broad}}$, and in the second signal sequence, narrow updating leads to a higher posterior belief than broad updating $\pi_t^{G,\text{narrow}} > \pi_t^{G,\text{broad}}$. Proposition 3 is a direct consequence of beliefs in narrow information processing not being invariant to the signal order.

Let $\theta_G = 0.7$, $\theta_B = 1 - \theta_G = 0.3$, $\lambda = 0.5$, and $\pi_0^G = \pi_0^B = 0.5$. Consider the sequences $S_3 = (g, b, g)$ and $S_3' = (b, g, b)$.

Case 1: $S_3 = (g, b, g)$

Narrow Information Processing $(n = 1)$**:**

Period 1 belief: $s_1 = g$

$$\mu_1^{G,\text{narrow}} = \frac{0.7 \cdot 0.5}{0.7 \cdot 0.5 + 0.3 \cdot 0.5} = 0.7 \tag{31}$$

$$\pi_1^{G,\text{narrow}} = 0.5 \cdot 0.5 + 0.5 \cdot 0.7 = 0.6 \tag{32}$$

Period 2 belief: $s_2 = b$

$$\mu_2^{G,\text{narrow}} = \frac{0.3 \cdot 0.6}{0.3 \cdot 0.6 + 0.7 \cdot 0.4} \approx 0.391 \tag{33}$$

$$\pi_2^{G,\text{narrow}} = 0.5 \cdot 0.6 + 0.5 \cdot 0.391 = 0.4955 \tag{34}$$

Period 3 belief: $s_3 = g$

$$\mu_3^{G,\text{narrow}} = \frac{0.7 \cdot 0.4955}{0.7 \cdot 0.4955 + 0.3 \cdot (0.5045} \approx 0.6962 \tag{35}$$

$$\pi_3^{G,\text{narrow}} = 0.5 \cdot 0.4955 + 0.5 \cdot 0.6962 = 0.59585 \tag{36}$$

Broad Information Processing ($n = 3$):

Since the batch is processed all at once from the prior $\pi_0^G = 0.5$, we can immediately calculate the final posterior belief.

Period 3 belief: $S_3 = (g, b, g)$

Batch likelihoods:

$$P(g, b, g \mid G) = 0.7^2 \cdot 0.3 = 0.147 \tag{37}$$

$$P(g, b, g \mid B) = 0.3^2 \cdot 0.7 = 0.063 \tag{38}$$

Then:

$$\mu_3^{(G,\text{broad})} = \frac{0.147 \cdot 0.5}{0.147 \cdot 0.5 + 0.063 \cdot 0.5} = 0.7 \tag{39}$$

$$\pi_3^{(G,\text{broad})} = 0.5 \cdot 0.5 + 0.5 \cdot 0.7 = 0.6 \tag{40}$$

Conclusion for Case 1: $\pi_3^{G,\text{narrow}} = 0.596 < 0.6 = \pi_3^{G,\text{broad}}$

Case 2: $S_3' = (b, g, b)$

Narrow Information Processing ($n = 1$):

Period 1 belief: $s_1 = b$

$$\mu_1^{G,\text{narrow}} = \frac{0.3 \cdot 0.5}{0.3 \cdot 0.5 + 0.7 \cdot 0.5} = 0.3 \tag{41}$$

$$\pi_1^{G,\text{narrow}} = 0.5 \cdot 0.5 + 0.5 \cdot 0.3 = 0.4 \tag{42}$$

Period 2 belief: $s_2 = g$

$$\mu_2^{G,\text{narrow}} = \frac{0.7 \cdot 0.4}{0.7 \cdot 0.4 + 0.3 \cdot 0.6} \approx 0.6087 \tag{43}$$

$$\pi_2^{G,\text{narrow}} = 0.5 \cdot 0.4 + 0.5 \cdot 0.6087 = 0.50435 \tag{44}$$

Period 3 belief: $s_3 = b$

$$\mu_3^{G,\text{narrow}} = \frac{0.3 \cdot 0.50435}{0.3 \cdot 0.5043 + 0.7 \cdot 0.49565)} \approx 0.3037 \tag{45}$$

$$\pi_3^{G,\text{narrow}} = 0.5 \cdot 0.5043 + 0.5 \cdot 0.3037 = 0.404 \tag{46}$$

Broad Information Processing ($n = 3$):

As before, we again immediately calculate the final posterior belief.

Period 3 belief: $S_3' = (g, b, g)$

Batch likelihoods:

$$P(b, g, b \mid G) = 0.3^2 \cdot 0.7 = 0.063 \tag{47}$$

$$P(b, g, b \mid B) = 0.7^2 \cdot 0.3 = 0.147 \tag{48}$$

Then:

$$\mu_3^{(G,\text{broad})} = \frac{0.063 \cdot 0.5}{0.063 \cdot 0.5 + 0.147 \cdot 0.5} = 0.3 \tag{49}$$

$$\pi_3^{(G,\text{broad})} = 0.5 \cdot 0.5 + 0.5 \cdot 0.3 = 0.4 \tag{50}$$

Conclusion for Case 2: $\pi_3^{G,\text{narrow}} = 0.404 > 0.4 = \pi_3^{G,\text{broad}}$

The example above shows that both $\pi_t^{G,\text{narrow}} < \pi_t^{G,\text{broad}}$ and $\pi_t^{G,\text{narrow}} > \pi_t^{G,\text{broad}}$ can hold depending on which information signals have been observed. Note that generally results for different sequences of information signals need not be monotonic in λ and can even reverse across λ levels. The interaction between λ and the type of information processing (narrow vs. broad) is non-trivial.

A.4 Proof of Proposition 4

Proposition 4: For $\lambda \neq 1$, narrow information processing leads to a higher variance in posterior beliefs than broad information processing.

Proof: To define the order of signals, let σ be a permutation of the set $\Sigma = \{1, 2, \ldots, t\}$, and define the permuted signal sequence as $S'_t = (s_{\sigma(1)}, \ldots, s_{\sigma(t)})$ with $S'_t \neq S_t$.

Broad Information Processing:

From Proposition A.2 we know that broad belief updating is invariant to the order of signals for each λ as long as the ordering of batches of information does not change. Formally, we have shown that because $P(S_{\sigma(t)} \mid \theta_j)$ only depends on the number of signals, not their order, for all permutations σ, $\mu_t^{j,\text{broad}}$ is the same. Hence, $\pi_t^{j,\text{broad}}$ is invariant to permutations.

Since $\pi_t^{j,\text{broad}}$ is invariant to permutations, it directly follows that conditional on the batch size and, for all σ:

$$\text{Var}_{\sigma\in\Sigma}\left[\pi_t^{j,\text{broad}}\right] = 0. \tag{51}$$

The intuition is straightforward. Holding the number of g and b signals constant for a given λ and n, each signal ordering would lead to the same posterior belief and the variance in posterior beliefs is zero.

Narrow Information Processing:

A similar reasoning applies to the case of narrow belief updating. In Proposition A.2 we have shown that narrow belief updating is not invariant to the order of signals as long as $\lambda \neq 1$ (i.e., the updating is non-Bayesian). Intuitively, even though all possible permutations start with the same prior belief, the second period belief already depends on the first observed signal and thus is subject to ordering. This divergence continues and amplifies over time due to the compounding of the non-linear updates. Formally, for two different permutations

$\sigma, \sigma' \in \Sigma$ with the corresponding signal history S_t and S_t', we generally have:

$$\pi_t^{j,\text{narrow}}(S_t) \neq \pi_t^{j,\text{narrow}}(S_t'). \tag{52}$$

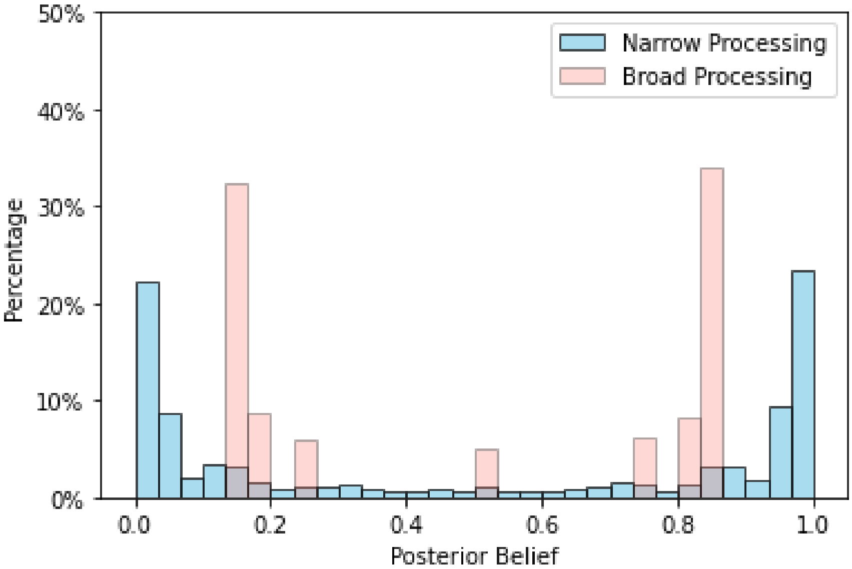
Since narrow updating yields a distribution of posterior beliefs across permutations, it follows that

$$\text{Var}_{\sigma \in \Sigma}\left[\pi_t^{j,\text{narrow}}\right] > 0, \tag{53}$$

which concludes the proof.

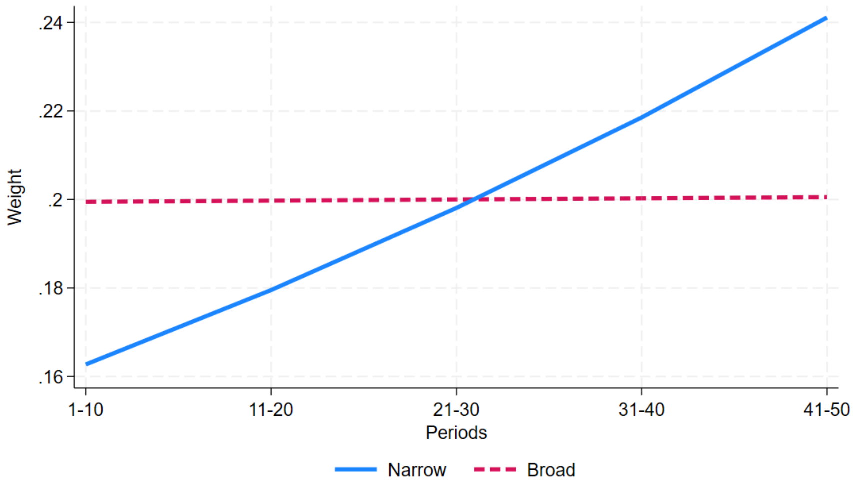
B Additional Figures and Tables

Figure B1: Framework Simulation



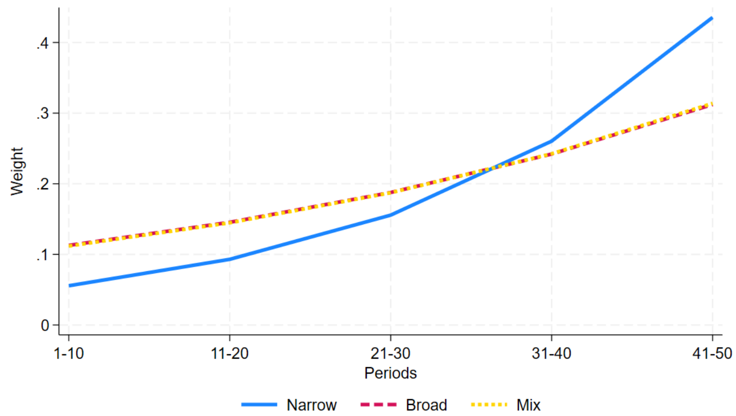
Note: This figure plots the distribution of posterior beliefs according to the narrow and broad updating protocol laid out in Section 2 using the following parameter values for 10,000 simulation runs: $t = 16$, $\theta_G = 0.7$, $\theta_G = 0.3$, $\lambda = 0.7$, $n = 1$ for narrow and $n = 16$ for broad.

Figure B2: Simulation Experiment 1: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 9 and simulated beliefs according to the narrow and broad updating protocol laid out in Section 2 using the parameter values of Experiment 1 for 10,000 simulation runs: $t = 50$, $\theta_i = s_i \in \{0.20, 0.21, ..., 0.80\}$, $\lambda = 0.7$, $n = 10$ for narrow and $n = 50$ for broad.

Figure B3: Intermediate Updates: Information Weights



Note: This figure plots the weights assigned to the information brackets based on the estimates of Equation 9 additionally accounting for the mixed treatment in the Finance Application.

Table B1: Experiment 1: Beliefs - Frequency of Price Increases

	(1)	(2)	(3)	(4)
all50 × *narrow*	-18.12***			
	(2.59)			
last10 × *narrow*	15.56***			
	(1.86)			
first40 × *narrow*		-14.50***		
		(2.07)		
last10 × *narrow*		11.94***		
		(1.51)		
first30 × *narrow*			-24.17***	
			(3.13)	
last20 × *narrow*			18.31***	
			(2.51)	
first20 × *narrow*				-4.55**
				(1.91)
middle20 × *narrow*				-9.95***
				(1.77)
last10 × *narrow*				11.79***
				(1.50)
FE	Yes	Yes	Yes	Yes
N	3700	3700	3700	3700
R^2	0.83	0.83	0.82	0.83

Note: This table shows regressions with the posterior belief in period 50 as dependent variable and the narrow information dummy, the frequency of price increases over different blocks of information, and their interaction with the narrow information dummy as independent variables. For conciseness, only the coefficients on the interactions are reported. For example, $all50_i$ corresponds to the frequency of price increases over all 50 periods, while $last10_i$ corresponds to the frequency of price increases over the last 10 periods. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B2: Experiment 2: Accuracy

	Overall	Overall	Negative	Positive
bayes	0.01	0.01	0.26***	-0.08**
	(0.01)	(0.01)	(0.04)	(0..04)
attention	-1.30***	-1.08***	-1.29***	-0.84**
	(0.28)	(0.28)	(0.33)	(0.38)
Controls	No	Yes	Yes	Yes
N	3,412	3,360	1,677	1,683
R^2	0.01	0.03	0.05	0.03

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, an attention dummy, which equals 1 if a subject is in the attention manipulation and 0 if they are in the narrow information treatment of the baseline experiment, and control variables as independent variables. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B3: Experiment 2: Information Weights

	(1)	(2)	(3)	(4)
δ_{broad}	0.97***	0.97***		
	(0.02)	(0.02)		
δ_{narrow}	-0.22***		0.74***	
	(0.02)		(0.02)	
$\delta_{attention}$	-0.01			0.96***
	(0.03)			(0.01)
a	48.52***	48.78***	48.44***	48.27***
	(0.13)	(0.20)	(0.23)	(0.23)
b	0.43***	0.43***	0.41***	5.45***
	(0.00)	(0.00)	(0.00)	(0.10)
N	5,108	1,696	2,004	1,408
R^2	0.83	0.86	0.79	0.86

Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B4: Applications: Information Weights

	Finance			Consumer Choice		
	(1)	(2)	(3)	(4)	(5)	(6)
δ_{broad}	0.77***	0.74***		1.01***	1.02***	
	(0.02)	(0.02)		(0.06)	(0.05)	
δ_{narrow}	-0.18***		0.62***	-0.29***		0.70***
	(0.03)		(0.02)	(0.07)		(0.05)
a	50.01***	51.30***	49.06***	0.21	-0.04	0.29
	(0.34)	(0.51)	(0.47)	(0.32)	(0.44)	(0.47)
b	0.57***	0.54***	0.60***	1.17***	1.29***	1.10***
	(0.01)	(0.01)	(0.00)	(0.11)	(0.14)	(0.15)
N	3,024	1,408	1,616	996	497	499
R^2	0.54	0.49	0.58	0.12	0.14	0.12

Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices in the finance application (Columns (1) to (3)) and on the weighted sum of the quality of ratings in the consumer choice application (Columns (4) to (6)) respectively. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B5: Experiment 1: Estimation Error – Number of Trials

Panel A: First Half	Overall	Overall	Negative	Positive
bayes	0.00	0.02	0.22***	-0.15**
	(0.01)	(0.01)	(0.05)	(0.06)
narrow	1.95***	1.91***	1.73***	2.00***
	(0.36)	(0.36)	(0.42)	(0.51)
Controls	No	Yes	Yes	Yes
N	1,850	1,828	931	897
R^2	0.02	0.04	0.07	0.05

Panel B: Second Half	Overall	Overall	Negative	Positive
bayes	-0.00	-0.00	0.27***	-0.11**
	(0.01)	(0.01)	(0.05)	(0.05)
narrow	1.45***	1.43***	1.31***	1.59***
	(0.34)	(0.33)	(0.40)	(0.43)
Controls	No	Yes	Yes	Yes
N	1,850	1,828	920	908
R^2	0.01	0.04	0.07	0.06

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, the narrow information dummy and control variables as independent variables for the first respectively second half of trials. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B6: Experiment 1: Information Weights – Number of Trials

Panel A: First Half			
	(1)	(2)	(3)
δ_{broad}	0.96***	0.96***	
	(0.03)	(0.04)	
δ_{narrow}	-0.27***		0.74***
	(0.04)		(0.02)
a	48.35***	48.51***	48.22***
	(0.23)	(0.30)	(0.34)
b	0.42***	0.43***	0.40***
	(0.00)	(0.01)	(0.01)
N	1,850	848	1,002
R^2	0.80	0.85	0.76
Panel B: Second Half			
	(1)	(2)	(3)
δ_{broad}	0.97***	0.98***	
	(0.03)	(0.02)	
δ_{narrow}	-0.19***		0.78***
	(0.03)		(0.02)
a	48.84***	49.06***	48.65***
	(0.20)	(0.27)	(0.30)
b	0.42***	0.43***	0.42***
	(0.00)	(0.01)	(0.01)
N	1,850	848	1,002
R^2	0.84	0.87	0.82

Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices in the first respectively second half of trials. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B7: Experiment 1: Estimation Error – Attentiveness

	Overall	Overall	Negative	Positive
bayes	-0.00	-0.00	0.23^{***}	-0.08^{*}
	(0.01)	(0.01)	(0.04)	(0.05)
narrow	1.86^{***}	1.76^{***}	1.66^{***}	1.83^{***}
	(0.34)	(0.33)	(0.39)	(0.43)
Controls	No	Yes	Yes	Yes
N	2,876	2,856	1,438	1,418
R^2	0.02	0.04	0.05	0.05

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, the narrow information dummy and control variables as independent variables, excluding subjects in the bottom quintile of the total working time distribution. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B8: Experiment 1: Information Weights – Attentiveness

	(1)	(2)	(3)
δ_{broad}	0.98***	0.98***	
	(0.03)	(0.02)	
δ_{narrow}	-0.25***		0.72***
	(0.03)		(0.02)
a	48.52***	48.81***	48.31***
	(0.18)	(0.24)	(0.25)
b	0.42***	0.43***	0.41***
	(0.00)	(0.01)	(0.01)
N	2,876	1,176	1,700
R^2	0.82	0.86	0.79

Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices, excluding subjects in the bottom quintile of the total working time distribution. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B9: Experiment 1: Estimation Error – Statistical Skill

Panel A: Low Skill				
	Overall	Overall	Negative	Positive
bayes	0.00	0.00	0.33***	-0.16***
	(0.01)	(0.01)	(0.05)	(0.05)
narrow	1.49***	1.31***	1.00**	1.54***
	(0.35)	(0.35)	(0.40)	(0.49)
Controls	No	Yes	Yes	Yes
N	2,284	2,252	1,126	1,126
R^2	0.01	0.02	0.06	0.04
Panel B: High Skill				
	Overall	Overall	Negative	Positive
bayes	-0.01	-0.01	0.11*	-0.08
	(0.01)	(0.01)	(0.06)	(0.06)
narrow	2.33***	2.40***	2.44***	2.36***
	(0.48)	(0.48)	(0.56)	(0.43)
Controls	No	Yes	Yes	Yes
N	1,416	1,404	725	679
R^2	0.04	0.05	0.07	0.04

Note: This table shows regressions with the estimation error (absolute difference between reported beliefs and Bayesian beliefs in period 50) as dependent variable and the Bayesian posterior, the narrow information dummy and control variables as independent variables for subjects with low respectively high self-reported statistical skill. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table B10: Experiment 1: Information Weights – Statistical Skill

Panel A: Low Skill			
	(1)	(2)	(3)
δ_{broad}	0.94***	0.94***	
	(0.03)	(0.02)	
$_{narrow}$	-0.25***		0.68***
	(0.03)		(0.02)
a	48.37***	48.39***	48.35***
	(0.20)	(0.27)	(0.30)
b	0.41***	0.42***	0.40***
	(0.00)	(0.01)	(0.01)
N	2,284	1,112	1,172
R^2	0.81	0.84	0.78
Panel B: High Skill			
	(1)	(2)	(3)
λ_{broad}	1.02***	1.01***	
	(0.03)	(0.03)	
λ_{narrow}	-0.19***		0.81***
	(0.03)		(0.03)
a	48.98***	49.51***	48.62***
	(0.24)	(0.29)	(0.34)
b	0.43***	0.44***	0.42***
	(0.00)	(0.01)	(0.01)
N	1,416	584	832
R^2	0.84	0.90	0.80

Note: This table shows nonlinear least squares regressions of the posterior belief in period 50 on the weighted sum of past changes in prices for subjects with low respectively high self-reported statistical skill. Standard errors are clustered at the individual level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

C Experimental Instructions and Screenshots

Instructions

General Setting

In period 0 the stock price of a fictional stock amounts to \$400. The stock price increases or decreases every period over 50 periods. The size of the price change is always \$10, either up or down. The likelihood of a *price increase* is the same for these 50 periods and is randomly determined in period 0. It can be any percentage number between 20% and 80%. Since there are equally many percentage numbers above and below 50%, the average probability of a price increase is 50%.

But if, for example, 62% is drawn, the likelihood of a price increase is 62% in each period and the likelihood of a price decrease is 38% (100%-62%) in each period. As such, price increases and decreases are indicative of the drawn likelihood of a *price increase* for the fictional stock.

Task

You will observe the price changes of the fictional stock over 50 periods. From time to time you are asked to estimate the randomly determined likelihood of a *price increase* for this stock. In particular, you have to enter an integer percentage number between 20% and 80%. The entire task is repeated up to 8 times for independent fictional stocks, i.e. each stock has its own randomly determined likelihood of a price increase.

On the next page the compensation scheme is described.

Compensation

In addition to the participation fee of £1.50, you can earn a bonus payment in the estimation task.

Three of your estimates are randomly selected at the end of the study. Your compensation increases by £0.30 for each estimate which is within 5% of the correct statistical probability of a price increase (e.g. the correct probability is 50% and your estimate is between 45% and

55%).

If you feel that you understand the instructions, press "Next" to proceed to answer a few comprehension questions before the experiment starts.

Comprehension Questions

Below we report the comprehension questions that subjects had to answer correctly after reading the instructions to proceed to the estimation task. Correct responses are displayed in bold.

1. You observe a price change of $-10, how do you have to update your probability estimate of a price increase?
    - I increase the probability estimate.
    - **I decrease the probability estimate.**
2. Assume the correct statistical probability of a price increase is 70%. Which probability estimate would be in the range such that you earn a bonus payment?
    - 55%
    - **67%**
    - 77%
    - 83%
3. Is a probability estimate of 50% reasonable before having seen any price changes?
    - **Yes**
    - No
    - Can't be answered

Screenshots of the Estimation Task

Figures C1 to C4 present the screens of the estimation task as seen by subjects in the experiment (using stock 1 as an example). One round consists of three sequential screens. First, subjects see the empty price-line chart, only indicating the starting price of 400 in period 0. Second, the price development appears on the price-line-chart. In the narrow information treatment, the price-line chart builds over time and subjects beliefs about the asset's fundamental are elicited every 10 periods. In the broad information treatment, subjects observe all price changes between period 1 and 50 at once and their beliefs are only elicited in period 50. Finally, subjects are informed that they reached the end of the estimation task for this stock and will continue with the next stock.

Figure C1: Start of the Estimation Task

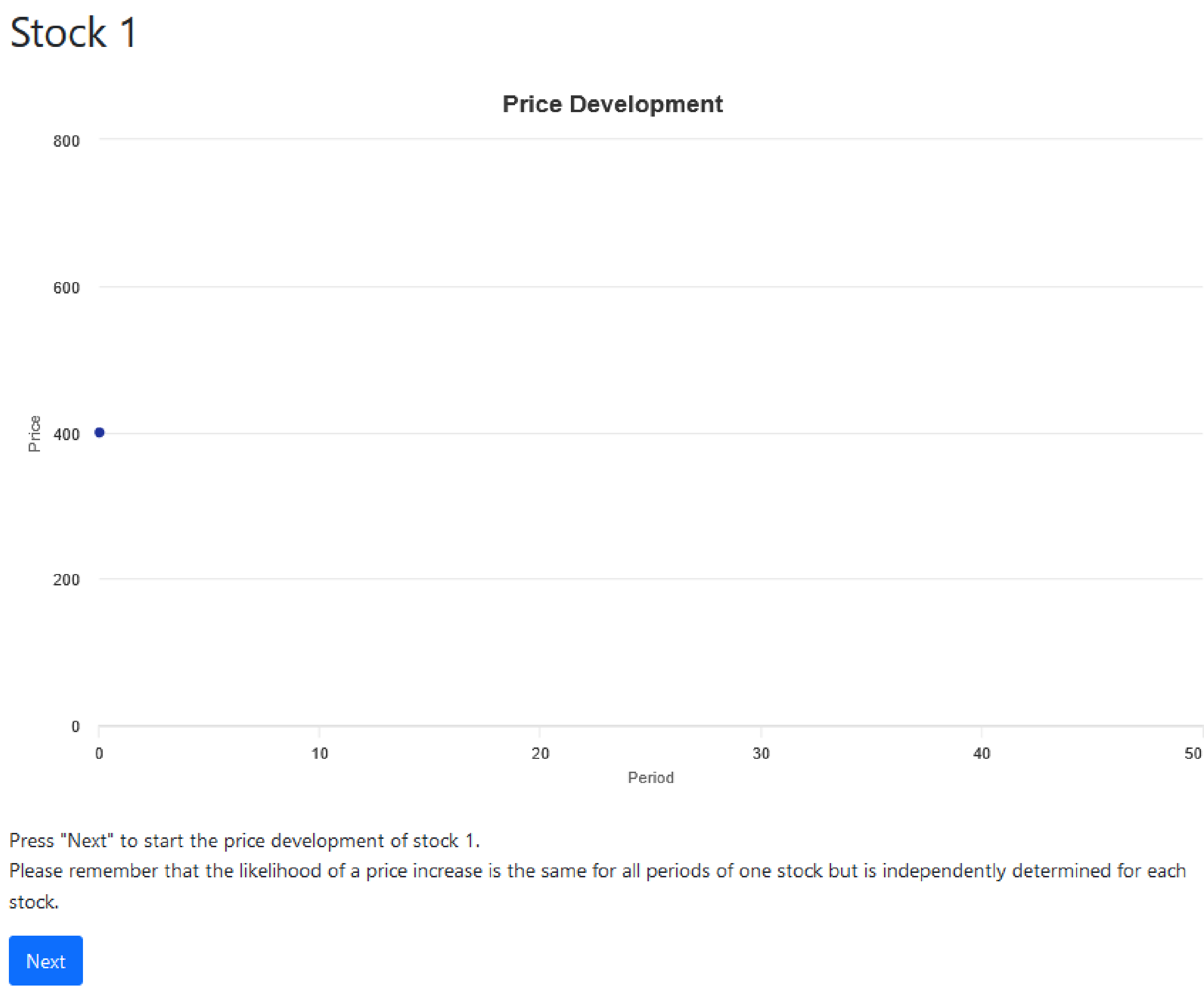
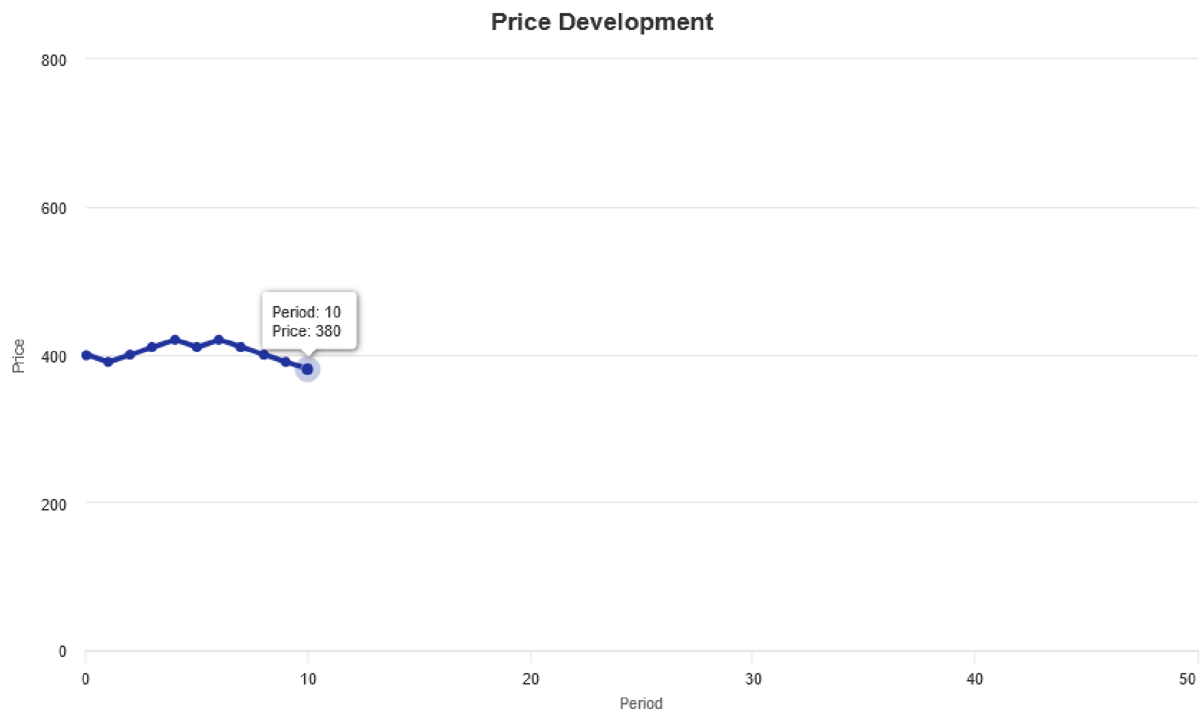


Figure C2: Belief Eliciation in Period 10

Stock 1



Given the price development, what do you think is the likelihood (in %) of a *price increase* for this stock? Please enter your estimate as an integer.

Next

Figure C3: Belief Eliciation in Period 50

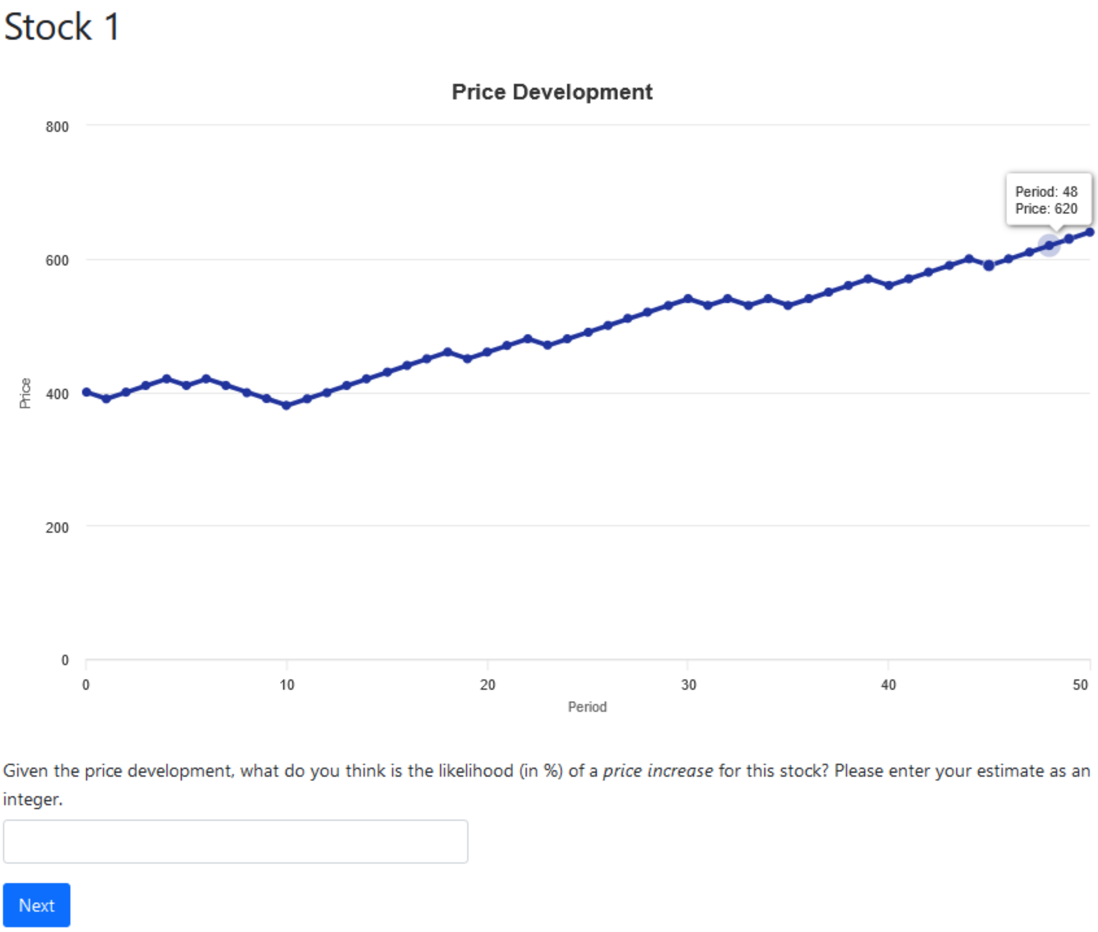
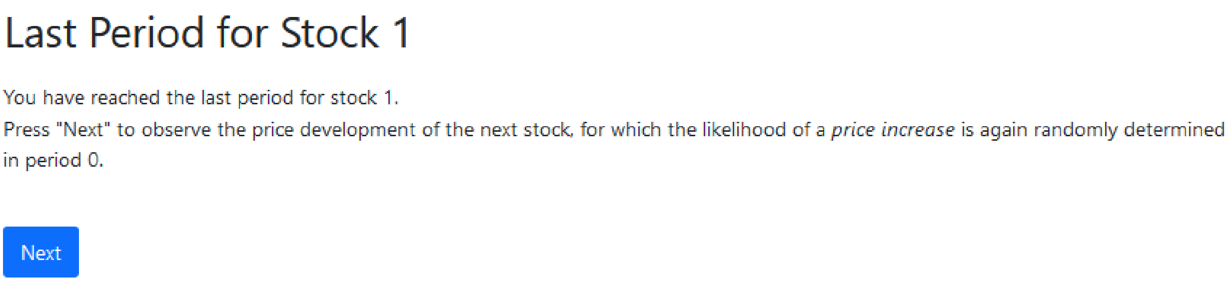


Figure C4: End of the Estimation Task



Screenshots of the Recall Task

Figures C5 to C8 present the screens of the recall task as seen by subjects in the experiment. The recall task consists of four sequential screens. First, the recall task is introduced to the subjects. Second, subjects are asked to recall the number of price increases and decreases they observed. Third, they are asked to recall the final price they observed. Finally, subjects are asked to recall the maximum number of subsequent price increases and decreases they observed.

Figure C5: Start of the Recall Task

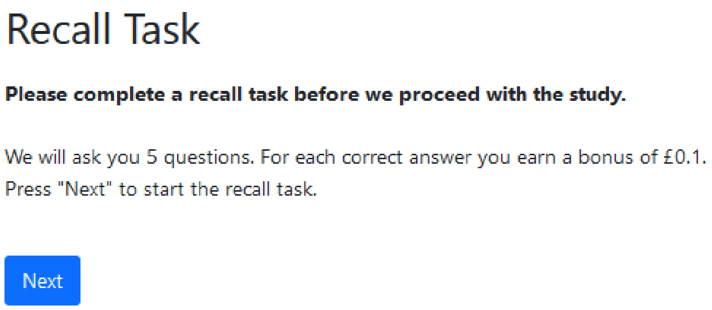


Figure C6: Recall Questions Page 1

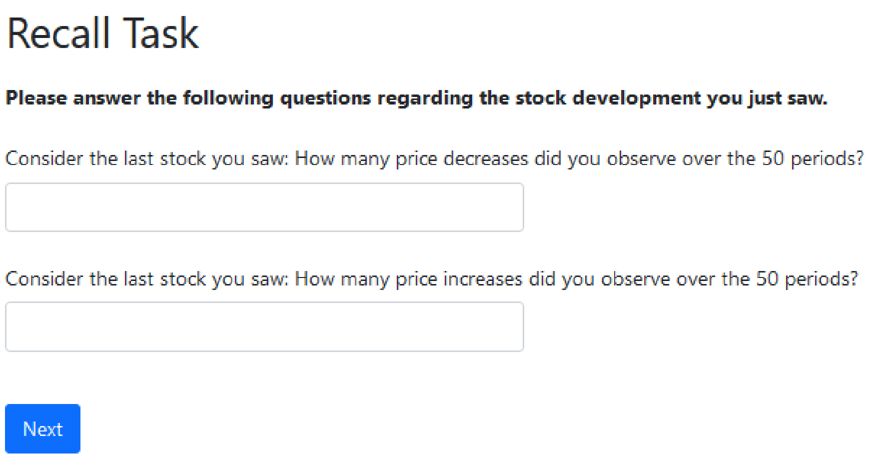


Figure C7: Recall Questions Page 2

Recall Task

Please answer the following questions regarding the stock development you just saw.

Consider the last stock you saw: What was the price of the stock in period 50?

Next

Figure C8: Recall Questions Page 3

Recall Task

Please answer the following questions regarding the stock development you just saw.

Consider the last stock you saw: If you had to guess, what is the maximum number of subsequent periods in which the the price repeatedly increased?

Consider the last stock you saw: If you had to guess, what is the maximum number of subsequent periods in which the the price repeatedly decreased?

Next

Screenshots of Attention Manipulation in Experiment 2

Figure C9 and Figure C10 present the screens of the attention manipulation as seen by subjects in the second experiment. The attention manipulation consists of two sequential screens. First, subjects start the rebuild of the price path. Second, once the rebuild is completed, subjects are asked to identify the asset's price for 5 periods.

Figure C9: Start of the Rebuild

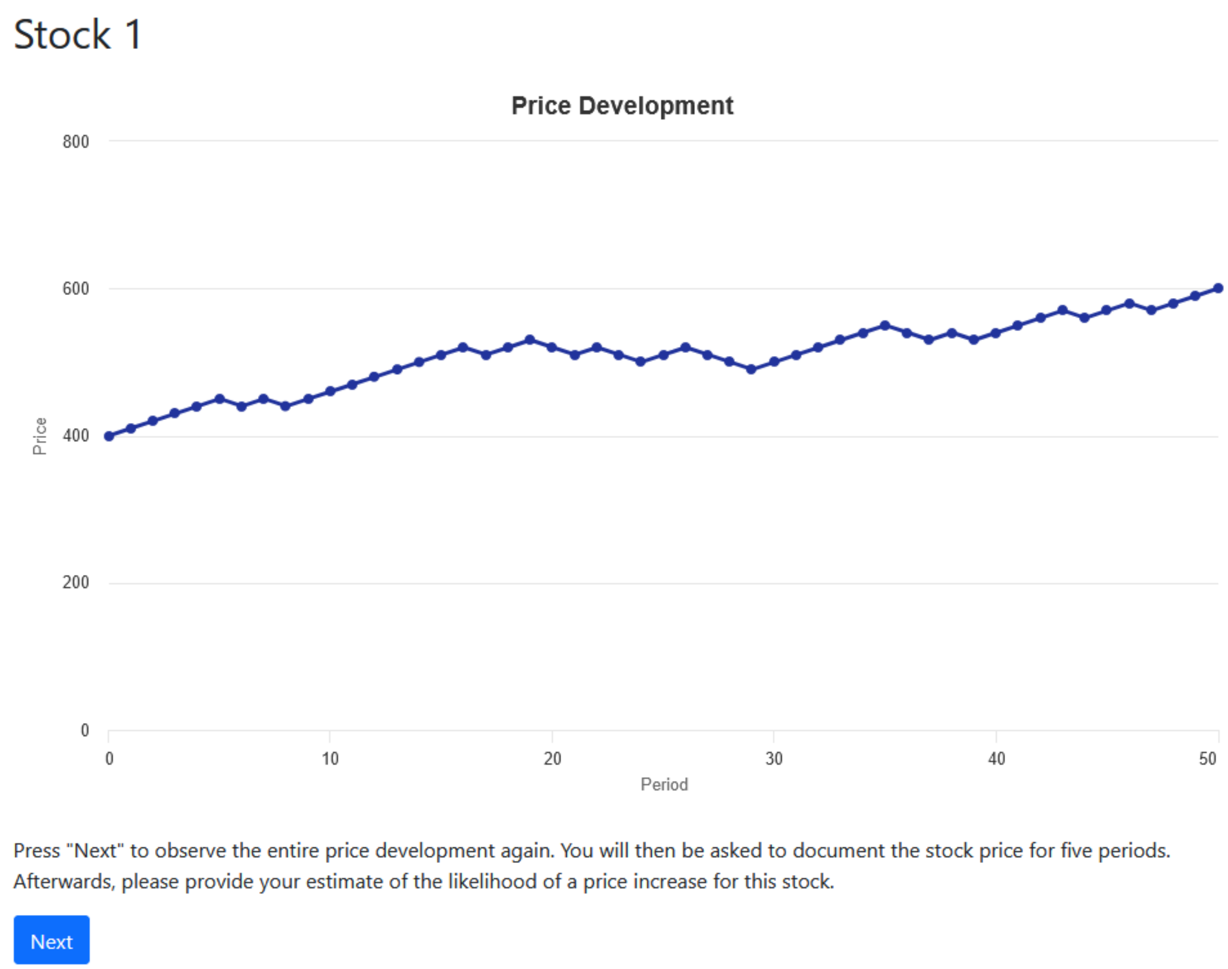
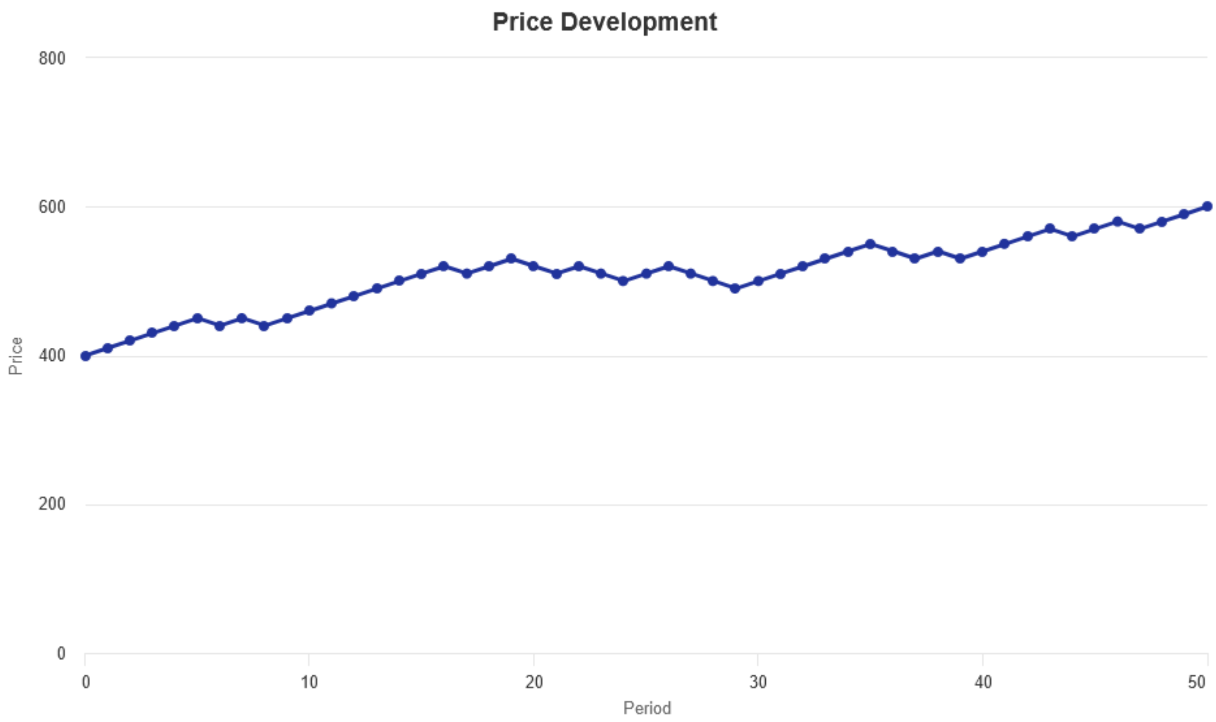


Figure C10: Price Identification

Stock 1



Please identify the stock price for the following five periods:

Period 21

Period 50

Period 34

Period 7

Period 16

Next

Screenshots of Product Ratings in Consumer Choice Application

Figure C11 and Figure C12 present the screens of the product ratings as seen by subjects in the consumer choice application. In the narrow information treatment, subjects observe 3 product ratings at a time with the new ratings being presented on top of the previous ratings. In the broad information treatment, subjects observe all 15 product ratings at once.

Figure C11: Product Ratings - Narrow Treatment

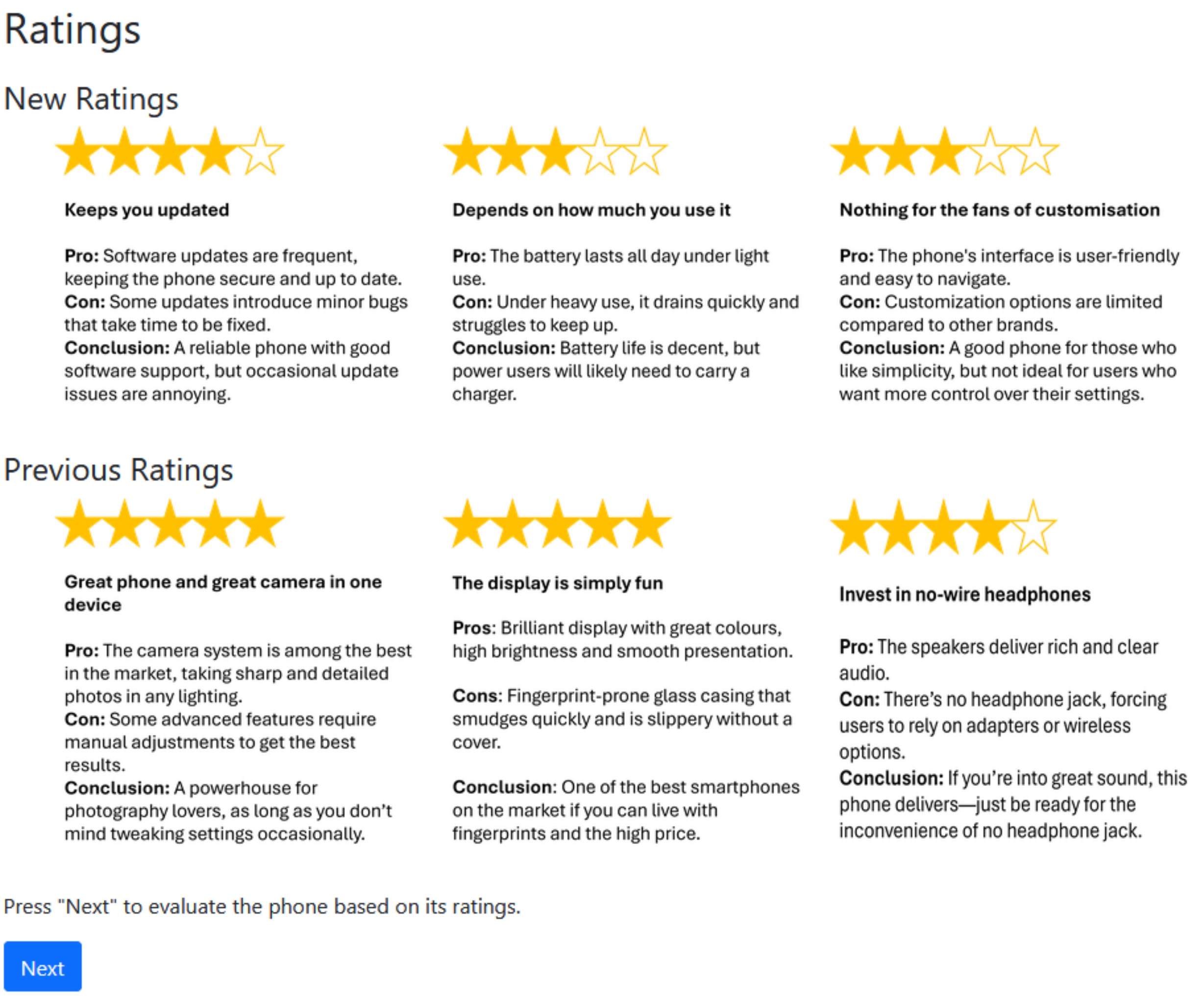


Figure C12: Product Ratings - Broad Treatment

Ratings

★★★☆☆

Nothing for the fans of customisation

Pro: The phone's interface is user-friendly and easy to navigate.
Con: Customization options are limited compared to other brands.
Conclusion: A good phone for those who like simplicity, but not ideal for users who want more control over their settings.

★☆☆☆☆

The screen is no fun

Pro: The screen is large and bright.
Con: The colors look unnatural, and the screen scratches easily despite claims of durability.
Conclusion: A big screen is great, but when it's fragile and poorly calibrated, it ruins the experience.

★★★★☆

Better be charged like normal

Pro: The phone charges very quickly.
Con: Wireless charging is slower than expected.
Conclusion: Fast charging is a great convenience, but wireless users may not get the same experience.

★★★☆☆

Rather good for offline-use

Pro: The phone supports 5G for faster internet speeds.
Con: It often loses signal in areas where other 5G phones still work fine.
Conclusion: While it has modern connectivity, reception issues make it unreliable.

★☆☆☆☆

Very unpleasant to have phone calls

Pro: The speakers are loud.
Con: The call quality is horrible, with muffled voices and constant background noise. Using headphones doesn't fix the issue.
Conclusion: A smartphone should, at the very least, work well for phone calls—this one fails at its most basic function.

★☆☆☆☆

No use

Pro: The phone was affordable.
Con: The performance is terrible—apps crash frequently, and it freezes during calls. It feels outdated despite being new.
Conclusion: Sometimes, a low price comes with too many compromises, making it not worth the purchase.

★★★★☆

Keeps you updated

Pro: Software updates are frequent, keeping the phone secure and up to date.
Con: Some updates introduce minor bugs that take time to be fixed.
Conclusion: A reliable phone with good software support, but occasional update issues are annoying.

★★★★★

I am very happy

Pro: The software is intuitive, smooth, and optimized perfectly for the hardware.
Con: Some customization options are still missing compared to other brands.
Conclusion: An incredibly polished experience that works right out of the box.

★★★★★

Definitely worth it

Pro: The performance is lightning-fast, even with demanding apps and games.
Con: The price is high, but you get what you pay for.
Conclusion: If you want top-tier speed and performance, this phone is worth the investment.

★★★☆☆

Rather heavy

Pro: The build quality feels solid and premium.
Con: It's heavier than expected, making one-handed use difficult.
Conclusion: A well-built phone, but the weight might be a dealbreaker for some users.

★★☆☆☆

The display is not it

Pro: The screen size is great for watching videos.
Con: The resolution is lower than expected, making text and images look pixelated.
Conclusion: A large screen is pointless if the display quality isn't sharp enough to match its size.

★★★☆☆

Depends on how much you use it

Pro: The battery lasts all day under light use.
Con: Under heavy use, it drains quickly and struggles to keep up.
Conclusion: Battery life is decent, but power users will likely need to carry a charger.

★☆☆☆☆

Simply impractical

Pro: The battery charges quickly.
Con: Unfortunately, it drains just as fast, barely lasting half a day. On top of that, the phone overheats easily.
Conclusion: A smartphone is useless if it can't last through a normal day without being constantly plugged in.

★★☆☆☆

Works down quite quickly

Pro: The phone has a nice, modern design.
Con: The materials feel cheap, and the back panel started coming loose after just a few months of use.
Conclusion: A smartphone should be durable, and this one doesn't seem built to last.

★☆☆☆☆

It keeps jerking

Pro: The design looks sleek at first glance.
Con: The phone lags constantly, even when performing simple tasks like texting or browsing the web. The touchscreen sometimes doesn't register inputs properly.
Conclusion: A smartphone should be fast and responsive, but this one is frustratingly slow and unreliable.

Press "Next" to evaluate the phone based on its ratings.

Next