Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Attention Constraints and Learning in Categories

Rahul Bhui, Peiran Jiao

MANAGEMENT SCIENCE
Vol. 69, No. 9, September 2023, pp. 5394–5404
ISSN 0025-1909 (print), ISSN 1526-5501 (online)
https://pubsonline.informs.org/journal/mnsc

Attention Constraints and Learning in Categories

Rahul Bhui,[a,*] Peiran Jiao[b]

[a] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142; [b] Department of Finance, Maastricht University, 6211 LM Maastricht, Netherlands

*Corresponding author

Contact: rbhui@mit.edu, https://orcid.org/0000-0002-6303-8837 (RB); p.jiao@maastrichtuniversity.nl, https://orcid.org/0000-0001-5207-0287 (PJ)

Received: September 23, 2022
Revised: January 25, 2023; March 21, 2023
Accepted: March 25, 2023
Published Online in Articles in Advance: May 30, 2023

https://doi.org/10.1287/mnsc.2023.4803

Abstract. Many decision makers are thought to economize on attention by processing information at the simpler level of a category. We directly test whether such category focus reflects an adaptive response to attention constraints, in five preregistered experiments using an information sampling paradigm with mouse tracking. Consistent with rational principles, participants focus more on category-level information when individual differences are small, when the category contains more members, and when time constraints are more severe. Participants are sensitive to the statistical structure of the category even when it must be learned from experience, and they respond to a latent shift in this structure. Beliefs about category members tend to cluster together more when category focus is high—a key element of rational inattention. However, this is counteracted by greater weight placed on salient and idiosyncratic information when the category is large. Our results broadly substantiate influential theories of categorical thinking, giving us a clearer view on the drivers and consequences of inattention.

History: Accepted by Marie Claire Villeval, behavioral economics and decision analysis.

Funding: This work was supported by the Office of Naval Research (N00014-21-1-2170), the Pershing Square Fund for Research on the Foundations of Human Behavior, and an NWO Vidi grant (VI.Vidi.201.059).

Supplemental Material: The data files and online appendix are available at https://doi.org/10.1287/mnsc.2023.4803.

Keywords: categorical thinking • rational inattention • information choice

1. Introduction

Our world is vast, but our attention is finite (Simon 1971, Kahneman 1973, Caplin 2016). We thus have to split our attention across the immense array of information available to us. In many situations, decision makers cope with this complexity by processing information at the simpler level of a category. For instance, investors have limited time and effort and cannot learn about all the countless stocks in a market. They might choose to study the value of an index rather than appraise each individual stock contained in that index (Peng and Xiong 2006). Similarly, managers may pay more attention to macroeconomic data than firm-level signals (Kacperczyk et al. 2016), analysts may compile aggregated rather than segmented information about branches of a company (Bens et al. 2018), and multiproduct firms may be more responsive to aggregate demand shocks than good-specific shocks (Pasten and Schoenle 2016).

Focusing on information at the category level leads agents to neglect heterogeneity among category members, resulting in economic anomalies. These anomalies are generally characterized by excess correlation in beliefs or outcomes of members in the same category and exaggerated differences across categories. For example, a range of inefficiencies has been tied to categorical investment patterns (aka "style investing") in behavioral finance (Barberis and Shleifer 2003), such as excess comovement of assets in the same class (Barberis et al. 2005). These phenomena can persist because attention may be naturally taxed more in thick markets with many firms, analysts, and investors, in contrast to the traditional view that distortions will be minimized in such markets.

Nevertheless, seminal theoretical work has shown how such category focus could be an individually efficient response when attention is scarce because category-level signals are informative about all category members, whereas idiosyncratic information pertains only to each member separately (Peng and Xiong 2006, Maćkowiak and Wiederholt 2009, Kacperczyk et al. 2016). This entails that attention to the category level versus the individual level should vary based on the costs and benefits induced by the information environment. These boundedly rational theories have provided influential explanations for the anomalous behavior of managers, firms, and households with large-scale economic consequences. However, despite the wide-ranging impact of these theories, we still lack direct evidence for the crucial assumption that people rationally balance attention to the category level and the individual level.

We conduct the most direct empirical test to date of rational inattention[1] applied to learning in categories, by developing a new laboratory paradigm. Our task was designed to transparently measure attention using mouse tracking while precisely controlling the statistical structure of information via an abstract sampling paradigm. Across five preregistered experiments, we test whether selective attention to category-level information adapts to the environment in line with rational principles.

People playing our "stock prediction game" had to accurately estimate the values of various hypothetical stocks based on a stream of incoming information.[2] These values were generated by a known categorical structure (following Peng and Xiong 2006). Participants were told that the stocks were all in the same industry and so the value of each was equal to the arithmetic sum of two latent components, a common industry-level factor (reflecting the category average) and a unique stock-specific factor (reflecting individual deviation from the average). These factors varied randomly and independently across periods.

In each period, participants could reveal noisy signals every half-second about any component (either the common industry factor or any one of the stock-specific factors) by hovering their mouse over the corresponding factor, until time ran out. They could only acquire signals for one factor at a time and therefore might have to alternate between factors depending on their strategy. Longer time spent on a factor meant more signals were acquired. Therefore, time was a proxy for attention, consistent with both theoretical tradition (Sims 2003) and high empirical correlation between the two (Caplin et al. 2020). This link is commonly made in process-tracing studies (Willemsen and Johnson 2011, Schulte-Mecklenbeck et al. 2017, Gabaix 2019) and enshrined in popular sequential sampling models of information processing (Krajbich 2019), as processing time is crucial in the brain's functioning (Pashler and Johnston 1998, Nobre and Coull 2010). In the field, viewing time of free, publicly available information is associated with reduced analyst error in earnings forecasts (Gibbons et al. 2021). We thus investigated theoretical predictions by measuring the amount of time spent mousing over each factor, and evaluating how this changed when we manipulated properties of the environment.

Existing theory generates sharp predictions in our paradigm. If attention is rationally deployed, one's category focus should adjust flexibly based on the value of information attainable at each level. Several implications follow from this central idea (Peng and Xiong 2006). First, when members of a category are similar, there is little to be gained by learning each one's unique qualities; second, category-level information reduces uncertainty about every member, and so its value scales with category size; third, continuing to accumulate information about a given variable yields diminishing returns. Thus, people should focus more on information at the category level when idiosyncratic variation is low relative to shared variation, when the category contains many members, and when attention constraints are severe. These hypothesized effects parallel important empirical phenomena. For example, managers may narrow their focus on aggregate information when market-wide volatility increases (Peng et al. 2007, Kacperczyk et al. 2016), when their firm sells many products (Pasten and Schoenle 2016), and when external events divert their attention (Huang et al. 2019).

Our experiments were designed to test these predictions. We found that people preferentially attended to information at the industry level when stock-specific variation was relatively lower (Experiments 1 and 5), when the industry contained more stocks (Experiment 2), and when time constraints were more severe (Experiment 3). There was no apparent effect of cognitive load caused by forcing signals to be kept in working memory (Experiment 4), which was consistent with the theory under the experimental parameters. People were sensitive to the prior variation at each level even when they were not given explicit information on these statistics or feedback on the accuracy of their predictions but had to learn them purely from the acquired signals; they were also able to adjust their attention allocation following a latent shift in the statistical structure of the category (Experiment 5).

We observed further signatures of rational inattention in our data. When a person's category focus is higher, their predictions of stock values in a given period should be more similar to each other because almost no differentiating information is being processed. Consistent with this, predictions tended to be less dispersed in periods with higher category focus—a key behavioral implication of models based on inattention (Peng and Xiong 2006, Kacperczyk et al. 2016). Moreover, category focus had a U-shaped relationship with prediction error that broadly matched the theoretical predictions in each experiment. Too much attention to the category leads to a detrimental neglect of individual differences, whereas too little attention means the category information is not being efficiently exploited. The intermediate allocation that balanced these opposing forces varied depending on the structure of the environment as described previously.

Categorical attention, behavior, and performance were thus linked to each other and the environment in accordance with rational inattention, with one exception. In conditions where attention is more strained, value predictions should be less dispersed (controlling for category focus) because less individuating information is acquired. However, when there were many stocks in Experiment 2, predictions were more dispersed rather than less. Model fitting revealed that in this condition, participants placed even more weight on the stock-specific signals, and on the most recent signals that were salient because their values were highlighted and displayed numerically (Bordalo et al. 2022). This finding suggests that as the task becomes more challenging, people may be more inclined to fixate on salient information, which can counteract the clustering of values that stems from categorical thinking.

Overall, we found that people adapted their degree of categorical focus broadly in line with rational principles. Our results substantiate core elements of influential theories of categorical information processing, while revealing how judgments might deviate from this benchmark. This work sharpens the link between categorical attention, behavior, and performance, giving us a clearer view on the drivers and consequences of inattention, and offering a reproducible platform for further investigations.

2. Rational Inattention and Learning in Categories

Our results speak to many prominent applications of rational inattention built around the hypothesized cognitive mechanism. Peng and Xiong (2006) theoretically demonstrate how the optimal allocation of attention can lead investors to focus on category-level information. Combining attention allocation with portfolio allocation, their model recapitulates several elements of style investing and empirical features of asset returns, such as excess comovement of assets within a category (Barberis and Shleifer 2003). Kacperczyk et al. (2016) theoretically and empirically analyze the performance of mutual fund managers and propose that an important part of manager skill involves properly balancing attention to macroeconomic aggregates and idiosyncratic firm-level data. They argue that in recessions (characterized by high aggregate volatility and price of risk), attention to aggregates should increase and fund outperformance should rise. Maćkowiak and Wiederholt (2009) posit that firms rationally attend more to idiosyncratic conditions when they vary more than aggregate conditions, which could explain why prices respond rapidly to sector-specific shocks and slowly to monetary policy shocks. Maćkowiak and Wiederholt (2015) extend this analysis to include households, with analogous implications for consumption patterns. These applications all rest on the assumption that decision makers rationally adapt their category focus.

Although past empirical work has aimed to tease out implications of rational inattention in market contexts (Peng et al. 2007, Drake et al. 2017, Huang et al. 2019, Choi and Gupta-Mukherjee 2022, Ehrmann and Jansen 2022, Liu et al. 2023), field settings pose many challenges to researchers. Naturalistic information structures are often opaque and high-dimensional, categorization schemes and information processing capacity may vary widely among individuals and circumstances, the lack of controlled experimental variation makes it hard to establish what causes attention to shift, and attention itself is difficult to clearly measure. Analyses of field data thus require various indirect and assumption-laden methods to infer critically important variables like attention (Gabaix 2019). Our experimental approach enables more straightforward tests.

Our research also relates to influential perspectives in cognitive science which maintain that cognitive resources like attention are spent where they have maximum benefit (Gottlieb 2018, Bhui et al. 2021, Summerfield and Parpart 2022). The results demonstrate how an adaptive response to limited processing capacity can yield systematic deviations from unconstrained Bayesian benchmarks. Past work has shown that when tasked with forming guesses of unknown feature values, people are thought to concentrate on prototypical category information, and some have argued this focus depends on how the information is to be used (Rehder et al. 2009, Braunlich and Love 2022). The rational predictions we test provide precise insight into conditions which stimulate attention to prototypical information and have escaped experimental scrutiny thus far.

3. Experiments 1–4

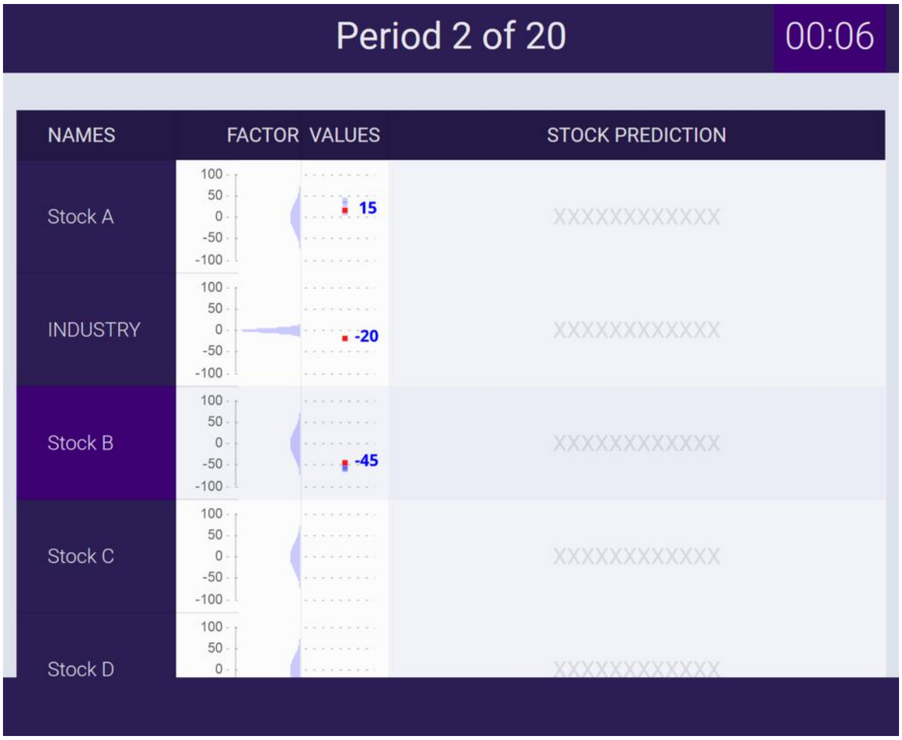
3.1. Participants

Five hundred eighty-four participants from the United States were recruited on Amazon Mechanical Turk, split across four experiments (Experiment 1, $n = 147$; Experiment 2, $n = 145$; Experiment 3, $n = 146$; Experiment 4, $n = 146$). They were paid a base of \$2 plus a bonus of up to \$6 that depended on performance.[3]

3.2. Procedure

All experiments used the same “stock prediction game” paradigm pictured in Figure 1, with some variations. In each period, participants had to estimate the values of several hypothetical stocks after selectively acquiring a stream of information about the components of value (a common industry-level factor and idiosyncratic stock-specific factors). The stocks were

Figure 1. (Color online) Screenshot of Experiment



Notes. Participants had to predict the values of hypothetical stocks, which were given by the sum of a common industry factor and idiosyncratic stock-specific factors. Noisy signals could be acquired moment-to-moment by mousing over any factor until time ran out.

abstractly labeled A, B, C, and so on. Values were given by the arithmetic sum of the two factors, meaning each stock's value was equal to the common industry-level factor plus an idiosyncratic stock-specific factor. These factors were generated independently in every period from zero-mean Gaussian distributions[4] portrayed on screen by sideways bell curves (so participants were shown these priors). Hence, the periods were effectively repetitions of the same estimation problem. All stock-specific factors had the same prior distribution to focus on the tension between the industry and stock levels rather than differences between stocks. In short, letting $c \sim \mathcal{N}(0, \sigma^2_{industry})$ be the industry factor and m_A, m_B, and $m_C \sim \mathcal{N}(0, \sigma^2_{stock})$ be stock-specific factors in a given period, participants had to make their best guess as to the stock values $v_A = c + m_A$, $v_B = c + m_B$, and $v_C = c + m_C$.

However, the exact factor values were not explicitly provided, and participants instead had to learn about them by mousing over the corresponding factor. While their mouse cursor was positioned over a given factor, a noisy Gaussian signal of its true value would be revealed every 500 ms, drawn accordingly from $\mathcal{N}(c, \sigma^2_{signal})$ or $\mathcal{N}(m_i, \sigma^2_{signal})$. Participants could mouse over any factor they wanted at any moment before a limited budget of time ran out, which was represented by an on-screen timer.

After time expired in a period, participants recorded their point prediction of every stock's total value (the sum of its two relevant factors) using a set of sliders. They did not provide predictions for the industry factor or any stock-specific factors by themselves. This phase had no time limit. Upon submitting these predictions, they were shown the true stock values and the magnitudes of their errors (only for feedback, as these did not affect any random variables in subsequent periods). At the end of the experiment, they were paid a bonus based on the mean squared error of their predictions in each period according to a quadratic loss function of which they were informed.[5]

We used a within-subjects design in each of the four experiments. All experiments consisted of two blocks of 10 periods each. Within each block the design parameters were fixed, and the treatments occurred across blocks. We implemented the following treatments, which were expected to increase the relative attention paid to the industry factor:

- Experiment 1's (variance) treatment increased the relative prior variance of the industry factor (simultaneously increasing $\sigma_{industry}$ from 5 to 30 and decreasing σ_{stock} from 30 to 5 to keep total variance constant).
- Experiment 2's (size) treatment increased the number of stocks (raising n_{stocks} from two to eight).
- Experiment 3's (time) treatment decreased the time budget (reducing available time from 20 to 8 seconds).

Table 1. Experimental Design Parameters

Experiment	Condition	$\sigma_{industry}$	σ_{stock}	σ_{signal}	n_{stocks}	Time (s)	Vanish
Experiment 1 (variance)	Low category variance	5	30	10	5	12	No
	High category variance	30	5	10	5	12	No
Experiment 2 (size)	Few category members	30	5	10	2	12	No
	Many category members	30	5	10	8	12	No
Experiment 3 (time)	Long time limit	30	5	10	5	20	No
	Short time limit	30	5	10	5	8	No
Experiment 4 (memory)	Signals remain	20	20	10	5	12	No
	Signals vanish	20	20	10	5	12	Yes
Experiment 5 (latent)	Low category variance	1	30	10	5	12.75	Yes
	High category variance	30	1	10	5	12.75	Yes
	Equal category variance	15	15	10	5	12.75	Yes

Note. In Experiment 5, the low and high category variance conditions occurred in the first half of periods, and the equal category variance condition occurred in the second half.

• Experiment 4's (memory) treatment increased the degree of cognitive load (rather than the signals remaining onscreen during the prediction stage, they vanished right as the next signal appeared).

The design parameters for each experiment are documented in Table 1 (along with those of Experiment 5, which will be described in a later section). The order of blocks was counterbalanced. The serial position of the industry factor on the screen was counterbalanced across subjects, but kept the same across blocks for any given subject. In Experiments 1–3, the last signal for each factor was highlighted and its number was displayed as pictured in Figure 1.

After reading the instructions, participants were provided with two self-paced practice periods in which they were told the true values of each factor and allowed unlimited time to sample information. This was intended to clearly explicate the task structure. They were subsequently asked two basic comprehension check questions to verify their understanding of the task (see the online appendix, Section EC.7.1).

3.3. Results

First, participants were able to perform reasonably well in the task. Predictions were moderately to highly correlated with true values, with median correlations ranging from 0.520 to 0.928 across experiments (see the online appendix, Section EC.4).

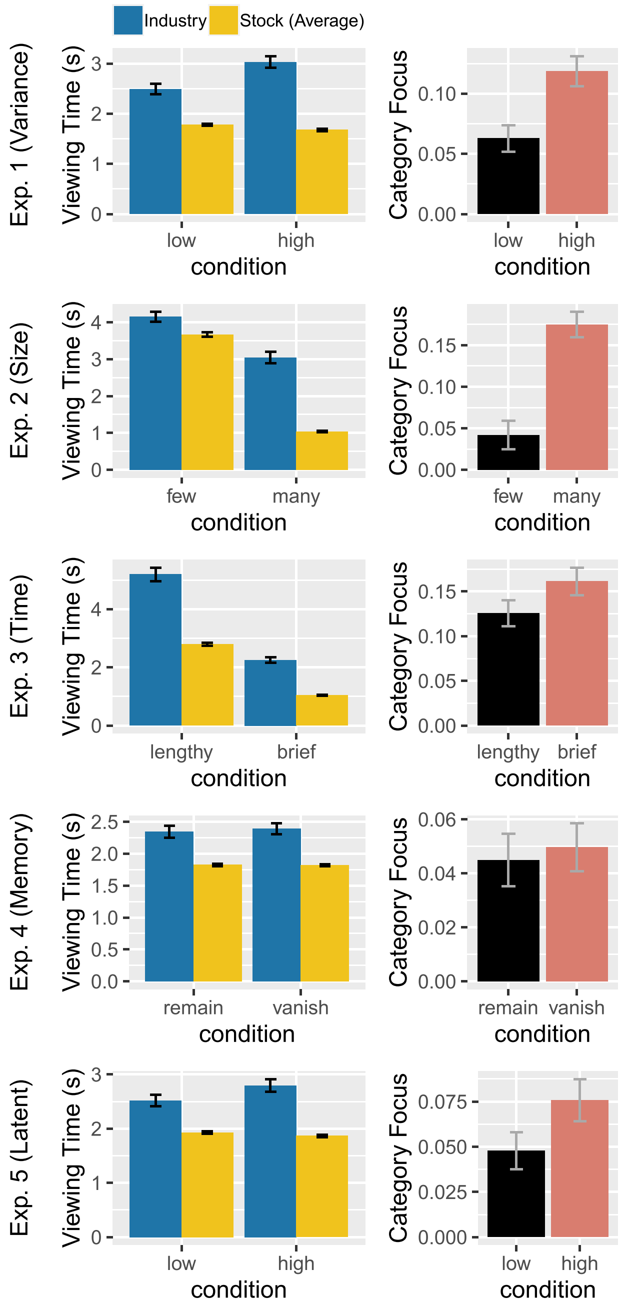
3.3.1. Category Focus. Our primary variable of interest is the category focus, which we define as the fraction of time spent mousing over the industry factor compared with the average stock-specific factor. For example, if out of the 12-second time limit, a participant spent 7 seconds attending to the industry factor and a total of 5 seconds attending to the five stock-specific factors (meaning an average of 1 second per stock-specific factor), the category focus would be $7/12 - 1/12 = 0.5$. This metric was used because it scales appropriately with the time limit and number of stocks and is motivated by the theoretical model we draw upon (see the online appendix, Section EC.1, for our streamlined Bayesian variant of the model in Peng and Xiong (2006)). If attention were merely split equally across all factors, the category focus would be fixed at zero in all conditions.[6]

The category focus is displayed for each experimental condition in Figure 2. The treatment effects appear to be in line with the first three predictions. These conclusions are formally supported by Bayesian random effects regressions reported in Table 2, which predict category focus based on the treatment condition, with subject-specific coefficients for both intercept and treatment effect. The regressions indicate positive effects of higher category-level variance ($P(\beta_{variance} > 0) > 0.999$), category size ($P(\beta_{size} > 0) > 0.999$), and time pressure ($P(\beta_{time} > 0) = 0.956$), but not of vanishing signals ($P(\beta_{load} > 0) = 0.679$). Our preregistered test criterion of $P(\beta > 0) > 0.95$ is met in the first three cases. Although the null effect from Experiment 4 ran counter to our initial expectations, as we will see later, it turns out to be consistent with the model due to the experiment's different design parameters. Thus, participants appear to alter their patterns of attention as predicted by the theory.[7]

3.3.2. Prediction Dispersion. We investigate another key behavioral signature of category thinking under rational inattention: When category focus is higher, predictions of stock values in a given period should be more similar to each other, because less individuating information is obtained.

The standard deviation[8] of participants' stock predictions is plotted conditional on the category focus in Figure 3. We also derive and plot the theoretical relationship between the two for comparison (see the online appendix, Section EC.1, for details), parameterized based on the experimental design.[9] The model implies that as category focus increases, the variance across stock predictions in a period should decline from the true prior variance of the stock-specific factor

Figure 2. (Color online) Attention Allocation Patterns



Notes. (Left) Mean time spent attending to each factor; (right) category focus (i.e., difference between proportions of time spent on category and average member). Low versus high relative category variance (Experiment 1 and first half of Experiment 5), few versus many category members (Experiment 2), lengthy versus brief time limit (Experiment 3), and signals remaining onscreen or vanishing (Experiment 4). Error bars depict 95% confidence intervals.

down to zero (plus any baseline response noise due, for instance, to the slider interface[10]). Consistent with this implication, predictions of stocks are generally more similar (i.e., their standard deviation is lower) when category focus is higher. The online appendix, Table EC.1, contains the results of Bayesian random effects regressions revealing this relationship.[11]

The theory also implies that prediction dispersion should be lower in conditions where attention is more strained, holding category focus constant. A violation of this is apparent in Experiment 2, however: Predictions are more variable, rather than less, when there are many stocks. We investigate this finding further using model fitting to capture how individual participants form predictions based on the signals they acquire (see the online appendix, Section EC.3). The model includes parameters to reflect the weight placed on industry signals, the weight placed on stock signals, the extra weight placed on the most recent signal from each factor, and response noise.[12] This analysis reveals that when there are many stocks (compared with when there are few), participants place even more weight on the stock-specific signals and on the most recent signals ($P(\beta > 0) = .990$ and 0.996 respectively, Bayesian signed rank test; see the online appendix, Table EC.5). The latter are salient because their values are displayed numerically and highlighted in red (Bazley et al. 2021). Such overweighting of a few noisy signals for each factor makes the predictions more dispersed. Regression analysis confirms that these weight parameters are indeed positively associated with dispersion in the stock predictions ($P(\beta > 0) = .999$ and 0.994; see the online appendix, Tables EC.6 and EC.7). This result suggests that when the task is complex due to scale, people may be inclined to fixate on salient information, which can counteract the correlation in predictions that stems from categorical thinking.

3.3.3. Prediction Accuracy. Although category focus is monotonically related to prediction dispersion, it should have a curvilinear relationship with prediction error. Attending too much to the category leads the agent to neglect heterogeneity, whereas attending too little prevents the agent from efficiently drawing upon the category information. An intermediate level of category focus balances these considerations (with corner solutions obtaining in more extreme cases). The exact location of this optimum depends on the environmental structure as described earlier.

Error in the stock predictions is plotted against the category focus in Figure 4. In the online appendix, Table EC.2 contains the results of Bayesian random effects regressions capturing this relationship. We also derive and plot the theoretical relationship between the two for comparison (see the online appendix, Section EC.1, for details), parameterized based on the experimental design as before. Participants exhibit more error than the theoretical bound likely because perceptual limitations prevent them from extracting the entire information content of the signals. However, the shapes of the theoretical curves are broadly recapitulated in the data.

Table 2. Treatment Effects on Category Focus

	Coefficient	Category focus		
		Mean	95% CI	$P(\beta > 0)$
Experiment 1 (variance)	Intercept	0.069	[0.042, 0.098]	>0.999
	Higher category variance	0.053	[0.026, 0.079]	>0.999
Experiment 2 (size)	Intercept	0.038	[0.000, 0.076]	0.975
	Larger category size	0.131	[0.091, 0.174]	>0.999
Experiment 3 (time)	Intercept	0.129	[0.091, 0.169]	>0.999
	Shorter time limit	0.030	[−0.006, 0.066]	0.956
Experiment 4 (memory)	Intercept	0.046	[0.026, 0.064]	>0.999
	Data points vanish	0.004	[−0.013, 0.021]	0.679
Experiment 5 (latent)	Intercept	0.048	[0.036, 0.059]	>0.999
	Higher category variance	0.028	[0.012, 0.043]	>0.999

Notes. Posterior estimates from Bayesian random effects regressions predicting category focus from experimental condition. $P(\beta > 0)$ denotes the posterior probability that the coefficient is positive. First half of periods included for Experiment 5. CI, credible interval.

The empirical incentive structure of the task thus seems commensurate with the theory.

These plots reveal that in Experiment 4, the optimal level of category focus is scarcely affected by the change in signal precision. This null effect occurs largely because the industry- and stock-level prior variances were equal in that experiment, negating the benefit of focusing on the category. Hence, the theory entails that under the actual task conditions, no effect should be expected. This observation underscores the importance of formal models that predict how category focus emerges from a complex interaction between environmental variables.

4. Experiment 5

In the previous experiments, participants were shown the prior variances of the industry and stock-specific factors. However, in many natural settings, this internal category structure might be unknown and would have to be learned from experience. Moreover, it could even change over time without being explicitly signposted. Experiment 5 (latent) explores whether people can cope under these more challenging conditions in two ways. First, can people adjust their attention allocation when the category's internal statistics must be learned from minimal information? Second, does the degree of category focus adapt when these statistics surreptitiously change? This investigation probes the boundaries of the claim that rational inattention applies well in repeated situations, where agents might discover the optimal strategy through experience (Maćkowiak et al. 2023).

4.1. Participants

Two hundred ninety-nine participants from the United States were recruited on Amazon Mechanical Turk according to the same criteria as the previous experiments. They were paid a base of \$3 plus a bonus of up to \$6 based on performance in the same way as the previous tasks. The sample size was increased because the treatment (described later) was between subjects rather than within subjects.

4.2. Procedure

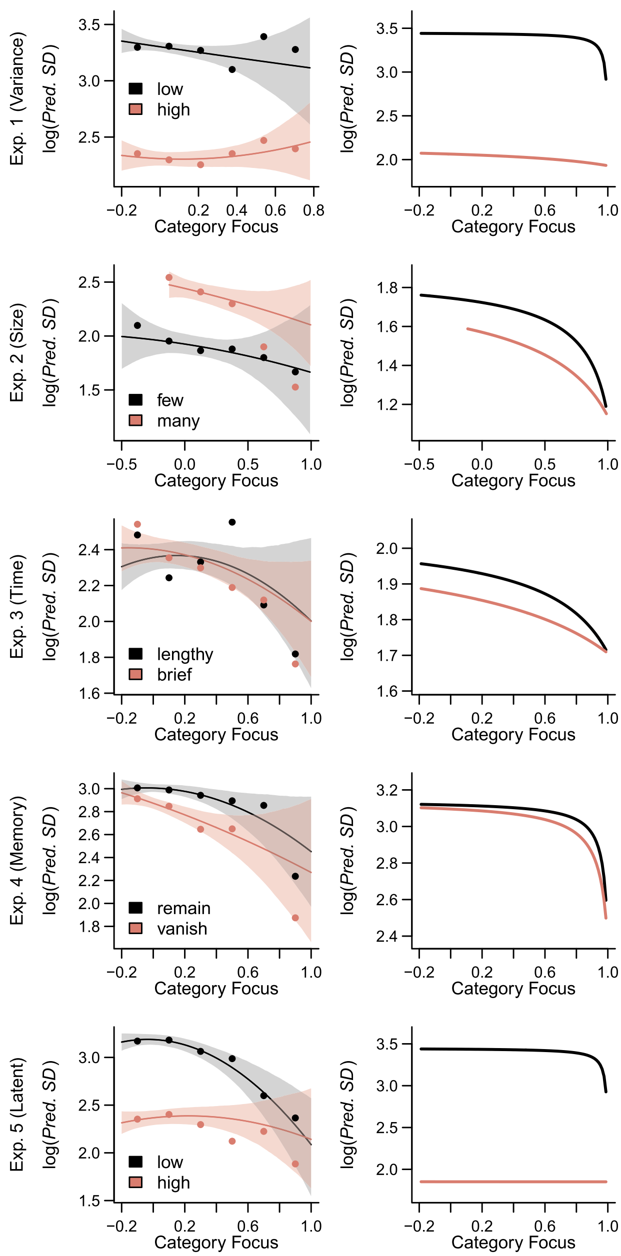
The basic elements of the task were similar to those of Experiment 1. However, in the first block of 10 periods, half of the participants experienced high relative category variance ($\sigma_{industry} = 30$, $\sigma_{stock} = 1$) and the other half experienced low relative category variance ($\sigma_{industry} = 1$, $\sigma_{stock} = 30$), again keeping total variance constant. In the second block of 10 periods, all participants encountered equal variances at both levels ($\sigma_{industry} = 15$, $\sigma_{stock} = 15$). The blocks were not explicitly demarcated, meaning there was no overt sign of this transition. Participants were also not told about the possibility of any changes. The design parameters are documented in Table 1.

4.3. Results

For consistency with the previous experiments, we conduct similar analyses and construct the same plots using the data from the first half of the periods in Experiment 5. The results are displayed in Figures 2–4. They replicate our results from Experiment 1, with small differences due to variation in the design parameters.

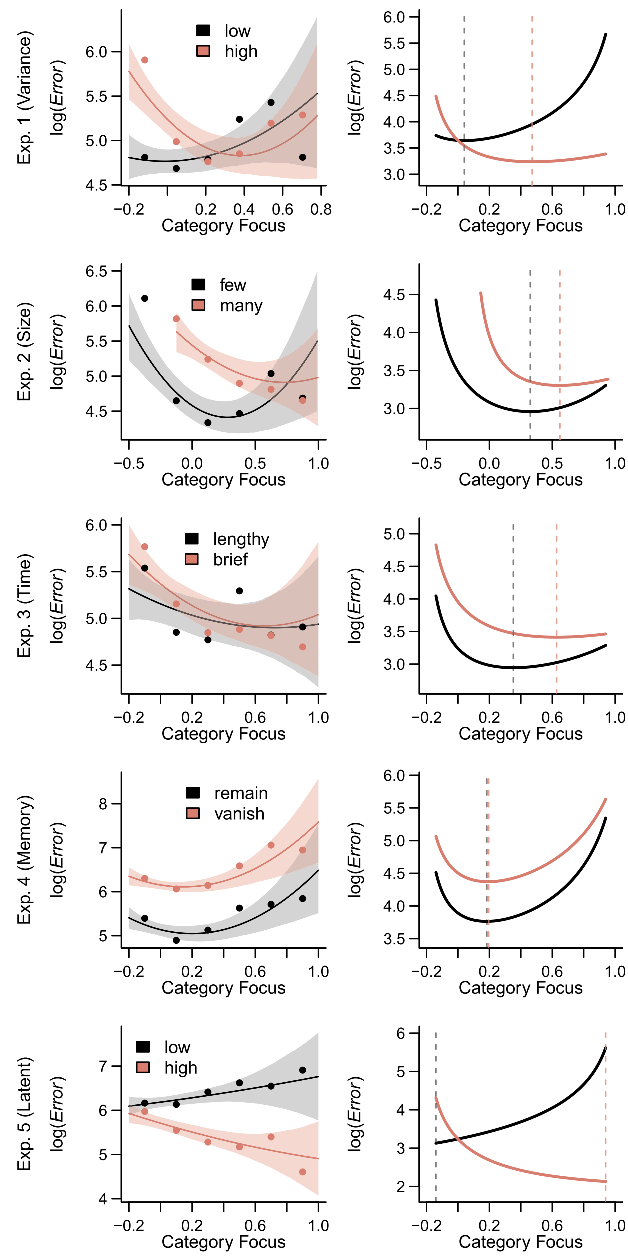
4.3.1. Learning Dynamics. The dynamics of category focus are displayed in Figure 5. Category focus begins at the same level in each condition, as there is hardly any way to determine the statistics with such little data. However, patterns of attention diverge across the first half of the task as environmental statistics are learned, consistent with rational principles. Category focus increases in the condition with high relative category variance where stocks have nearly identical values, whereas it decreases in the condition with low relative category variance where stocks have nearly uncorrelated values (interaction $P(\beta > 0) > .999$; see Bayesian random effects regressions reported in Table 3).

Figure 3. (Color online) Prediction Dispersion and Category Focus



Notes. The dependent variable log(Pred.SD) is the logarithm of one plus the standard deviation of stock value predictions in a given period. (Left) Data, binned averages with quadratic regression lines and 95% credible intervals from Bayesian random effects regressions. (Right) Theory, plus response noise.

Figure 4. (Color online) Prediction Error and Category Focus



Notes. The dependent variable log(Error) is the logarithm of one plus the mean squared error in stock value predictions in a given period. (Left) Data, binned averages with quadratic regression lines and 95% credible intervals from Bayesian random effects regressions. (Right) Theory, plus response noise; dashed lines indicate error-minimizing levels of category focus.

After the latent shift occurs halfway through the task, these trends change. Category focus starts to converge again (interaction $P(\beta < 0) = .976$), as both conditions then have the same environmental statistics.[13]

5. Concluding Remarks

Influential theories rely on the pivotal assumption that learning at the category level emerges from the rational allocation of limited attention (Peng and Xiong 2006, Maćkowiak and Wiederholt 2009, Kacperczyk et al. 2016). We conduct the most direct empirical test of this assumption to date. We develop an abstract sampling paradigm which lets us tightly control the structure of the information environment and reveal attention transparently using mouse tracking. These design features enable us to directly weigh the data against implications

Table 3. Regression Results: Dynamics of Category Focus in Experiment 5 (Latent)

	Category focus					
	First half			Second half		
Coefficient	Mean	95% CI	$P(\beta > 0)$	Mean	95% CI	$P(\beta > 0)$
Intercept	0.071	[0.043, 0.099]	(>0.999)	0.020	[−0.014, 0.053]	(.878)
High cond. (first)	−0.018	[−0.057, 0.022]	(.189)	0.085	[0.040, 0.130]	(>0.999)
Period	−0.005	[−0.010, −0.001]	(.009)	0.001	[−0.002, 0.004]	(.735)
High cond. (first) × *Period*	0.011	[0.005, 0.017]	(>0.999)	−0.004	[−0.008, 0.000]	(.024)

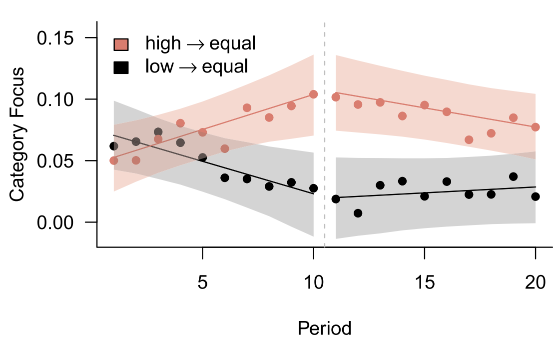
Notes. Posterior estimates from Bayesian random effects regressions with category focus regressed on experimental condition and period. The 95% credible intervals are in brackets and $P(\beta > 0)$ are in parentheses.

of rational inattention. The results indicate that people flexibly adjust their patterns of attention to category information broadly in line with rational principles.

Future research can build on our work in several ways. First, more complex interactions between attention and decision making should be experimentally characterized. Interdisciplinary research shows how attention can play an important role in many consequential settings, and some argue that a large part of behavioral economics reflects inattention (Gabaix 2019). We investigated the steps of information processing and expectation formation but not subsequent choice behavior. The objective function we used was meant to reflect a generic goal involving the accurate evaluation of each item. This allowed us to cleanly isolate the formation of beliefs. We expect our findings to be relevant for choice particularly in cases where judgment is directly related to decision making (Peng and Xiong 2006). It would nevertheless be valuable to explore how the downstream uses of information affect the manner in which information is processed upstream, as has begun to be systematically studied in neuroscience (Gottlieb 2018). Our task serves as a useful springboard for such studies.

Second, other tests will be needed to determine whether our findings extend to more covert forms of attention (Carrasco 2011). In contrast to some (but not all) models of rational inattention (Mondria 2010), participants could not shape information with full flexibility. This may be appropriate for settings where information is available in restricted forms, and might serve as a sensible approximation in other cases, but testing subtler aspects of attention would require another paradigm.

Figure 5. (Color online) Dynamics of Category Focus in Experiment 5 (Latent)



Notes. Dashed gray line indicates the change point in the variance structure. Regression lines shown from Bayesian random effects model with 95% credible intervals.

Finally, our results indicate the need for theories that incorporate multiple facets of attention. Two forms of attention are commonly considered in cognitive psychology: "top-down"—the volitional allocation of attention based on motivations and goals in a task; and "bottom-up"—the capture of attention by stimuli which are salient due to properties such as contrast, surprise, and prominence. We focused primarily on the former, by varying the usefulness of different signals. Our paradigm does allow elements of the latter to intrude, such as by the salience of recent or prominent signals, or the order in which factors are displayed on the screen. These might be explored further to test theories of attention guided by both task relevance and stimulus salience (Hefti and Heinke 2015). For example, Heinke (2019) extends a model of rational inattention to incorporate the visibility of signals, producing a rich interplay between active and passive information processing. Our findings suggest that the scale of the decision problem may influence attention via both pathways. A unified formal perspective on attention is essential for a truly complete framework.

Acknowledgments

The authors thank Florian Fröhlich for helpful research assistance.

Endnotes

[1] Throughout the paper, we use a broad definition of rational inattention which retains the core idea of selective attention that is adaptively deployed but does not restrict it to information-theoretic formalisms. Our task similarly places some limits on the shape of attainable signals, so that we can directly and transparently measure attention. This allows us to focus on—and is consistent with—our motivating applications to categorical information choice (Kacperczyk et al. 2016) rather than the most subtle internal properties of information processing.

[2] Minimizing the variance of beliefs can emerge naturally from a broader optimization problem, such as in the setting of Peng and Xiong (2006) where agents invest in the stock market under budget and attention constraints.

[3] Section EC.7 in the online appendix details our preregistration information and exclusion criteria for all five experiments.

[4] We use Gaussian prior distributions for concordance with the theories we draw on (Peng and Xiong 2006, Kacperczyk et al. 2016). Furthermore, Gaussian signals can be optimal in information-theoretic formulations under quadratic loss (Maćkowiak et al. 2023).

[5] The bonus was given by $\$6 - \sum_{t=1}^{20} \frac{1}{n_t} \sum_{i=1}^{n_t} (\hat{v}_{i,t} - v_{i,t})^2/200$, where t denotes the period, n_t denotes the number of stocks in period t, i is the stock index, and $\hat{v}_{i,t}$ is the prediction of stock i's value in period t, $v_{i,t}$. The penalty term was capped at $6/20 = $0.30 in each period so that the payment would not drop below $0. The objective function in our experiment was to minimize the mean squared error of predictions, whereas that of Peng and Xiong (2006) was to maximize expected lifetime utility. However, the two approaches share similar qualitative properties. More discussion and derivations can be found in the online appendix, Section EC.1.

[6] Category focus may be negative if less attention is paid to the industry than to the average stock-specific factor. It can range from $-1/n_{stocks}$ (no attention to industry) to one (all attention to industry).

[7] Due to the i.i.d. Gaussian nature of the signals, the theory does not constrain the dynamic sequence of information processing. Because we accordingly focus on the overall amount of attention paid to different information sources rather than the sequence, we report analyses of the attention trajectory in the online appendix, Section EC.5. In addition, serial position effects are depicted in Section EC.6.

[8] Because of high skewness, we log transform the standard deviation and later the error of the stock predictions.

[9] The plotted theoretical predictions for Experiment 4 assume that vanishing signals translates into cutting signal precision in half. Qualitative implications are not appreciably different with other fractions.

[10] When plotting the theoretical predictions, we roughly calibrate the level of response noise to the data. Specifically, we set the standard deviation of response noise to be the 10th percentile of prediction standard deviations across all included periods in each experiment. This provides a balance between finding the minimal level of variation (corresponding to irreducible noise) while avoiding undue influence from outliers (either participants or periods with unusually low noise).

[11] We use a quadratic regression specification for category focus to permit an analysis that does not assume the model is correct, but is flexible enough to allow for some nonlinearity which we expect based on the theory.

[12] The model captured individual judgments well, as model predictions were moderately to very highly correlated with participant responses, with median correlations ranging from 0.590 to 0.934 across experiments (see the online appendix, Section EC.4).

[13] Persistent changes do remain throughout the task. Category focus remains higher in the condition starting with higher category focus ($P(\beta > 0)$ greater than 0.95 for periods 7 onward, Bayesian t tests).

References

Barberis N, Shleifer A (2003) Style investing. *J. Financial Econom.* 68(2):161–199.

Barberis N, Shleifer A, Wurgler J (2005) Comovement. *J. Financial Econom.* 75(2):283–317.

Bazley WJ, Cronqvist H, Mormann M (2021) Visual finance: The pervasive effects of red on investor behavior. *Management Sci.* 67(9):5616–5641.

Bens DA, Monahan SJ, Steele LB (2018) The effect of aggregation of accounting information via segment reporting on accounting conservatism. *Eur. Accounting Rev.* 27(2):237–262.

Bhui R, Lai L, Gershman SJ (2021) Resource-rational decision making. *Current Opinion Behav. Sci.* 41:15–21.

Bordalo P, Gennaioli N, Shleifer A (2022) Salience. *Annu. Rev. Econom.* 14(1):521–544.

Braunlich K, Love BC (2022) Bidirectional influences of information sampling and concept learning. *Psych. Rev.* 129(2):213–234.

Caplin A (2016) Measuring and modeling attention. *Annu. Rev. Econom.* 8:379–403.

Caplin A, Csaba D, Leahy J, Nov O (2020) Rational inattention, competitive supply, and psychometrics. *Quart. J. Econom.* 135(3):1681–1724.

Carrasco M (2011) Visual attention: The past 25 years. *Vision Res.* 51(13):1484–1525.

Choi HM, Gupta-Mukherjee S (2022) Analysts' reliance on industry-level vs. firm-specific information: Implications for information production. *J. Bank. Finance* 143:106555.

Drake MS, Jennings J, Roulstone DT, Thornock JR (2017) The comovement of investor attention. *Management Sci.* 63(9):2847–2867.

Ehrmann M, Jansen DJ (2022) Stock return comovement when investors are distracted: More, and more homogeneous. *J. Internat. Money Finance* 129:102742.

Gabaix X (2019) Behavioral inattention. *Handbook of Behavioral Economics: Applications and Foundations 1*, vol. 2 (Elsevier, New York), 261–343.

Gibbons B, Iliev P, Kalodimos J (2021) Analyst information acquisition via EDGAR. *Management Sci.* 67(2):769–793.

Gottlieb J (2018) Understanding active sampling strategies: Empirical approaches and implications for attention and decision research. *Cortex* 102:150–160.

Hefti A, Heinke S (2015) On the economics of superabundant information and scarce attention. *Œconom. History Methodology Philosophy* 5(1):37–76.

Heinke S (2019) Active and passive information acquisition: Application to financial markets. Preprint, submitted August 13, https://dx.doi.org/10.2139/ssrn.3391221.

Huang S, Huang Y, Lin TC (2019) Attention allocation and return co-movement: Evidence from repeated natural experiments. *J. Financial Econom.* 132(2):369–383.

Kacperczyk M, Van Nieuwerburgh S, Veldkamp L (2016) A rational theory of mutual funds' attention allocation. *Econometrica* 84(2):571–626.

Kahneman D (1973) *Attention and Effort*, vol. 1063 (Prentice-Hall, Englewood Cliffs, NJ).

Krajbich I (2019) Accounting for attention in sequential sampling models of decision making. *Current Opinion Psych.* 29:6–11.

Liu H, Peng L, Tang Y (2023) Retail attention, institutional attention. *J. Financial Quant. Anal.* 58(3):1005–1038.

Maćkowiak B, Wiederholt M (2009) Optimal sticky prices under rational inattention. *Amer. Econom. Rev.* 99(3):769–803.

Maćkowiak B, Wiederholt M (2015) Business cycle dynamics under rational inattention. *Rev. Econom. Stud.* 82(4):1502–1532.

Maćkowiak B, Matějka F, Wiederholt M (2023) Rational inattention: A review. *J. Econom. Literature* 61(1):226–273.

Mondria J (2010) Portfolio choice, attention allocation, and price comovement. *J. Econom. Theory* 145(5):1837–1864.

Nobre AC, Coull JT (2010) *Attention and Time* (Oxford University Press, New York).

Pashler H, Johnston JC (1998) Attentional limitations in dual-task performance. Pashler H, ed. *Attention* (Psychology Press, New York), 155–189.

Pasten E, Schoenle R (2016) Rational inattention, multi-product firms and the neutrality of money. *J. Monetary Econom.* 80:1–16.

Peng L, Xiong W (2006) Investor attention, overconfidence and category learning. *J. Financial Econom.* 80(3):563–602.

Peng L, Xiong W, Bollerslev T (2007) Investor attention and time-varying comovements. *Eur. Financial Management* 13(3):394–422.

Rehder B, Colner RM, Hoffman AB (2009) Feature inference learning and eyetracking. *J. Memory Language* 60(3):393–419.

Schulte-Mecklenbeck M, Johnson JG, Böckenholt U, Goldstein DG, Russo JE, Sullivan NJ, Willemsen MC (2017) Process-tracing methods in decision making: On growing up in the 70s. *Current Directives Psych. Sci.* 26(5):442–450.

Simon HA (1971) Designing organizations for an information-rich world. *Comput. Comm. Public Interest* 72:37.

Sims CA (2003) Implications of rational inattention. *J. Monetary Econom.* 50(3):665–690.

Summerfield C, Parpart P (2022) Normative principles for decision-making in natural environments. *Annu. Rev. Psych.* 73:53–77.

Willemsen MC, Johnson EJ (2011) Visiting the decision factory: Observing cognition with MouselabWEB and other information acquisition methods. Schulte-Mecklenbeck M, Kuehberger A, Johnson JG, eds. *A Handbook of Process Tracing Methods for Decision Research* (Psychology Press, New York), 21–42.